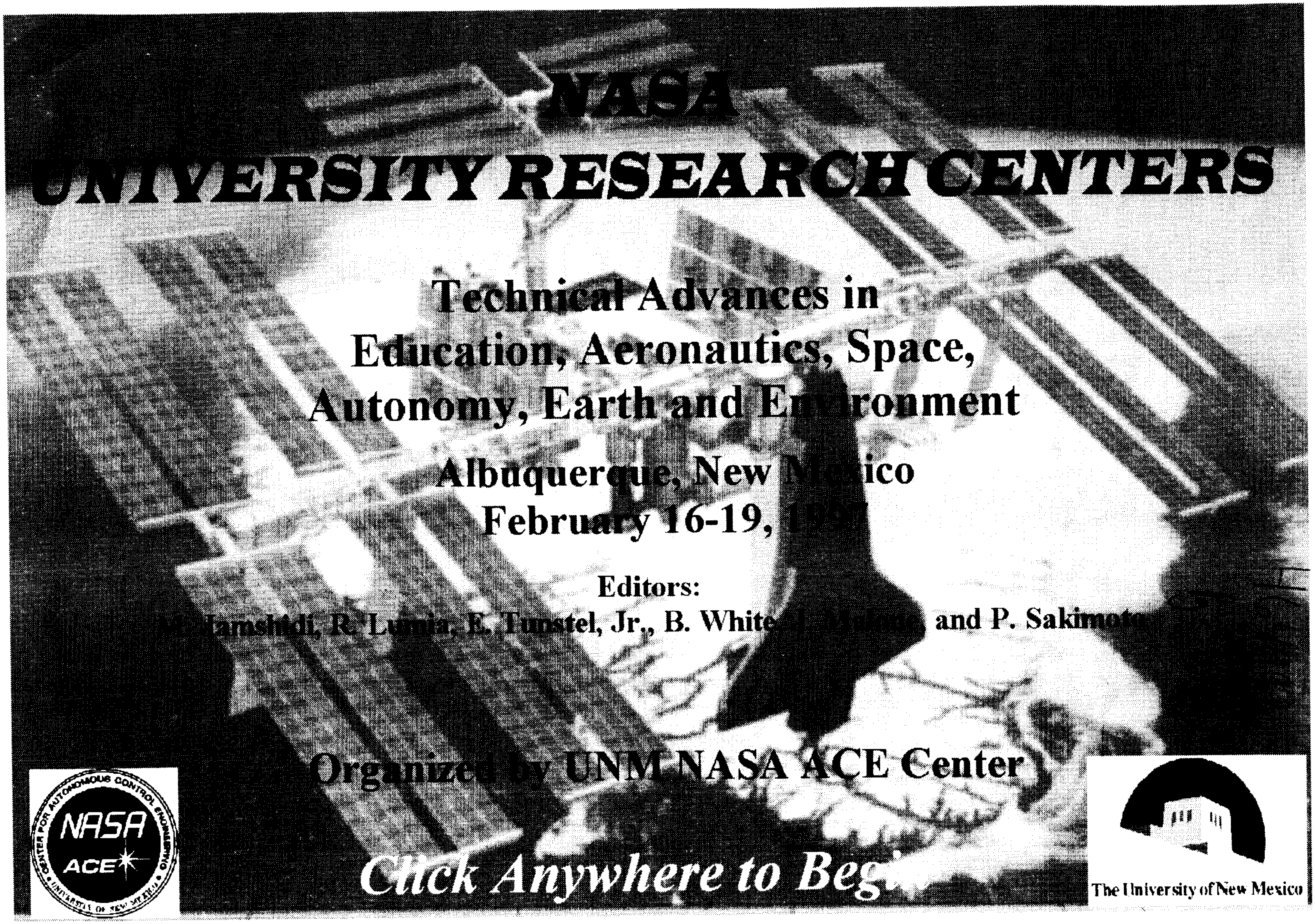

NASA
UNIVERSITY RESEARCH CENTERS
Technical Advances in
Education, Aeronautics, Space,
Autonomy, Earth and Environment
Albuquerque, New Mexico
February 16-19, 1997
Editors:
M. Jamshidi, R. Lumia, E. Tunstel, Jr., B. White, and P. Sakimoto
Organized by UNM NASA ACE Center
NASA ACE
Click Anywhere to Begin
The University of New Mexico

URC-TC'97

NASA University Research Centers - Technical Conference

Main Menu

Search By

URC-TC'97

NASA University Research Centers - Technical Conference

Author Index

A	B	C	D	E	F	G	H	I	J	K	L	M
N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Authors can be found by clicking the appropriate letter that corresponds to their last name. To view an author's paper, click the paper number next to their name. To return to this index, you must close that paper (File > Close)

Main Menu

A

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

H

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

L

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

O

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

X

Y

AUTHOR INDEX

URC-TC'97

NASA University Research Centers - Technical Conference

Subject Index

A	B	C	D	E	F	G	H	I	J	K	L	M
N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Subjects can be found by clicking the appropriate letter that corresponds to the subject. To view a paper, click the paper number next to the subject. To return to this index, you must close that paper (File > Close)

Main Menu

SUBJECT INDEX

A

B

SUBJECT INDEX

D

E

SUBJECT INDEX

F

H

SUBJECT INDEX

SUBJECT INDEX

L

SUBJECT INDEX

M

N

SUBJECT INDEX

O

P

SUBJECT INDEX

Q

SUBJECT INDEX

GOTO S - PAGE 2

T

SUBJECT INDEX

U

SUBJECT INDEX

URC-TC'97

NASA University Research Centers - Technical Conference

For more information about this CD and the conference, you may contact,

NASA ACE Center
Room 110
EECE Building
University of New Mexico
Albuquerque, New Mexico, 87131, U.S.A.

Phone: (505) 277-0300
Fax: (505) 277-4681
Email: jamshidi@unm.edu
adrienne@houdini.unm.edu
URL : http://ace.unm.edu

Main Menu

How This CD-ROM Works

Acrobat™ gives you many different ways to **view**, **display** and **scroll** the papers that are on a CD-ROM. The tool bar contains many different icons that will give you the ability to view a document any way you desire. See the "Description of Tool Bar Icons" in the main menu (they are also on the back side of your inside CD-ROM insert).

Hyperlinks (links): The color **BLUE** represents a hyperlink. By clicking on the blue areas, you will automatically be taken to that location.

Navigational Buttons: The [◀◀] [▶▶] buttons represent the previous and next pages and may have a caption describing the next area/section. The [**TOP**] button is a quick link to the "top" (page 1) of the current section you are in. These are not displayed in the Authors' papers.

Papers: When viewing <u>any</u> paper from on this CD-ROM, you <u>MUST</u> click on [File], then [Close] to exit that paper. This will take you back to the exact area from where you hyperlinked from. Any Graphic/photo with a blue box around it is a link to that image in a visually enhanced form.

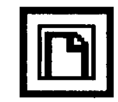
Searching the CD: See the About Acrobat™ Search in the Main Menu for a detailed guide on the search procedures and features.

Main Menu

Here are some suggestions from Omnipress: Screen Display

Tool Bar Icons

Display: When you select a paper to view, it is set up to "Fit Width." This will give you the maximum (best) viewing area of the page. If you change your viewing mode, you can return to the maximum viewing area by clicking the "Fit Width" tool bar icon.

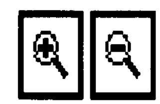


Fit Width

Mouse "Hand" Pointer: Use the "Hand" mouse pointer whenever possible. You can actually "grab" the screen by click-and-holding anywhere on the screen.

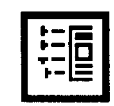
Hand

Zoom: Use the "Zoom + or -" tool bar icons to zoom in or out. Return quickly to the normal view with the "Fit Width" tool icon. Use the "zoom" tool icon to draw a box(click-and-drag) around the area you want zoomed.

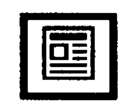


Zoom + / -

Bookmarks: When viewing certain indexes, bookmarks will appear on the left side of your monitor. These bookmarks provide a useful means of navigation. Each bookmark is an actual link to that area of the book. Bookmarks that have a *triangle* "▹" to their left have sub-bookmarks(expand by clicking on the triangle). Using the two tool bar icons, located to the extreme left of the tool bar, will allow you to switch quickly between "Bookmarks" display and "Page Only display.



Bookmarks



Page Only

General Notes

How to Print: To print the full paper or a page from any paper, select File > Print. A dialog box appears. Printing Range options are:

1. *All* - full document
2. *Current page* - prints the current page that is displayed.
3. *Pages:* _____ *to* ______ - this range allows you to print for examples, pages _5_ to _15_. The page numbers can be found on the information bar on the bottom-left of the screen and they do correspond to the actual page number of each paper.

Note: Acrobat documents may not print to some dot-matrix printers.

Printing Visually Enhanced Graphics: To capture the best print image of any graphic/photo, click the desired Graphic/photo, then print that page. Any graphic/photo not printed while in the visually enhanced form may have a poor image.

Exiting this CD-ROM: To exit this CD-ROM, you should close out all of the files. To do this, simply click (Window > Close all) from the Acrobat menu bar at the top of your screen. Then quit the Acrobat Reader 2.1 application(File > Exit).

TOP

Acrobat™ Search

Acrobat™ Search gives you <u>full-text search capabilities</u> for collections of Portable Document Format (PDF) documents that have been indexed. The added feature and benefits of Acrobat™ Search are included in this CD-ROM product.

Unlike a normal “Find” feature of a word processing program(performed by means of the binoculars only tool bar icon), which searches for words by reading every word on every page of a single paper, the added Acrobat Search feature searches full-text indexes. The full-text indexes created can represent hundreds or thousands of documents, and compared to searching a document word-by-word one page at a time—searching a full-text index is very fast. A full-text index is an alphabetized list of all the words and terms used in a collection of documents. Acrobat Search uses full-text indexes to quickly find words and terms in the documents without having to open the documents.

Search Query displays the Search window, which you can use to perform a new search or refine the current search.

Main Menu

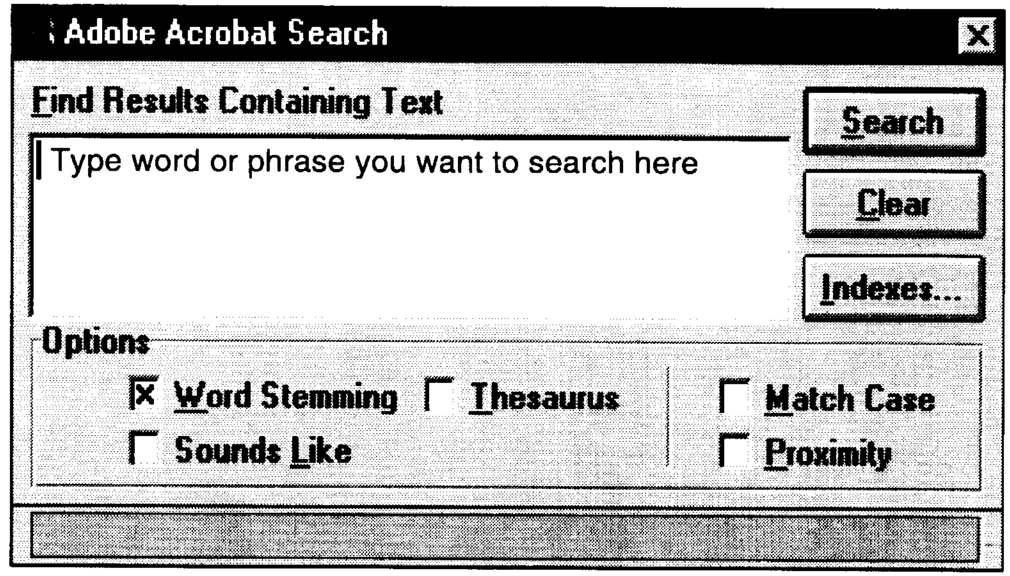
Search Results displays the Search Results window, which you can use to quickly open any found document.

Steps in Conducting a Search

1. To begin a search, simply type the word or phrase in the main white box.
(see Search Examples).

2. Select one of the many different options to use when you perform a search.
(see Search Options.)

3. Finally, click on Search
The Results window will automatically appear.



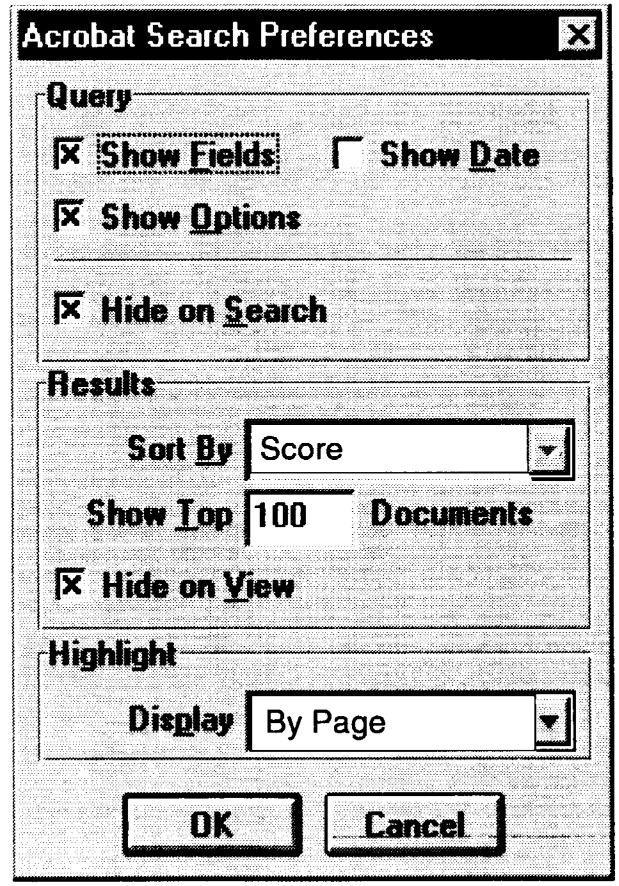
To Refine a Search

Performing a search on the original search results will better narrow the search results. Follow steps 1 & 2, then hold down the [Ctrl]key-Windows users, [options]key-Mac users, the search button becomes a Refine button. Click "Refine" to narrow your original search even more.

TOP

Search Preferences

By clicking EDIT > PREFERENCES > SEARCH... the Acrobat Search Preferences window appears. This will list the options that will affect the Search window and Search Results window.



How to set your preferences

Clicking "Show Fields" in the Query section will add the options of searching just the Title, Subect, Author, or Keywords. Thus, in addition to searching the the full text search index, you can search specific fields.

We suggest that you note any changes that you make in order to restore the default settings.

The next page(Acrobat Search Window) demonstrates the effects to the Acrobat Search window after "Show Fields" has been selected.

TOP

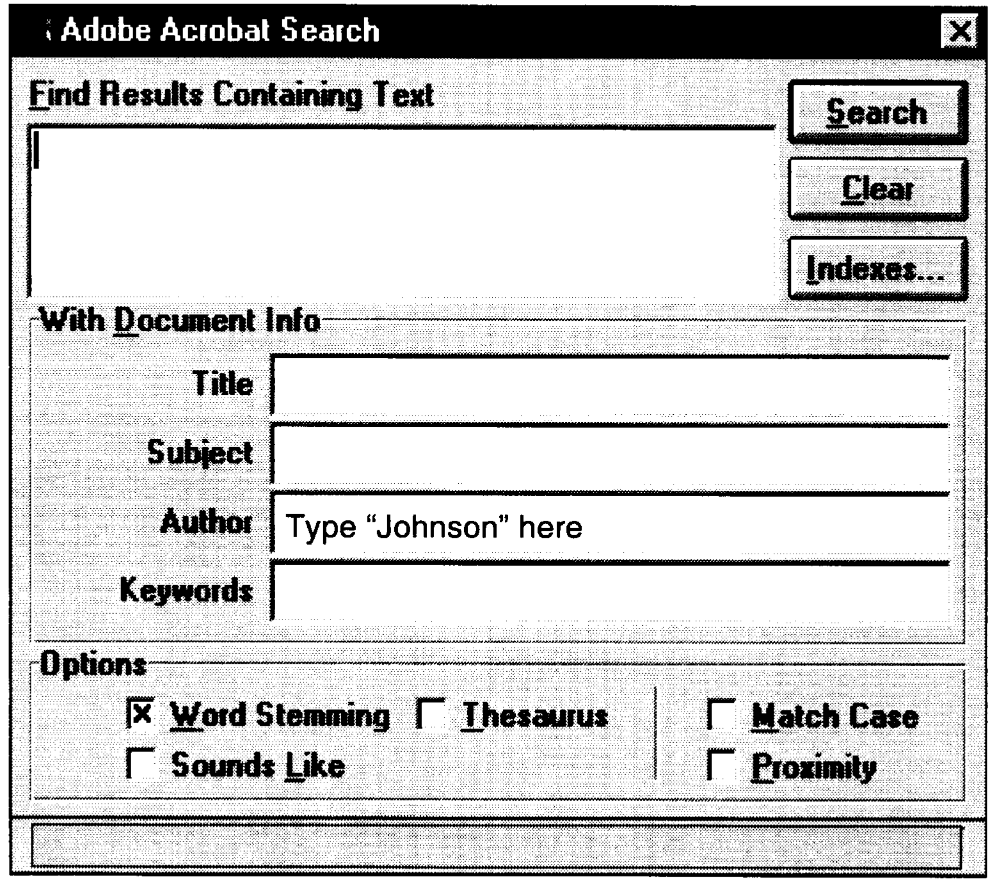
Acrobat Search Window

(after "Show Fields" in the Search Preferences has been changed)

The search fields Title, Subject, Author, and Keywords now appear within the Search window.

Searching by any one of these fields is now possible.

For example, if you wanted to search for an author that had the last name of "Johnson", you would type "Johnson" into the Author field(instead of the main white box) and click Search. The results would appear for only authors that had the name of "Johnson".



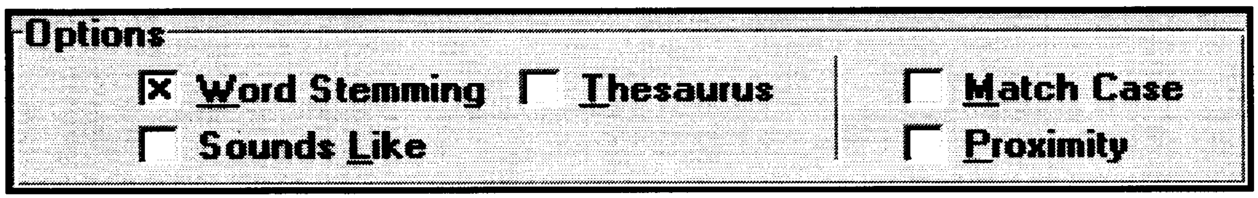
Search Examples

Query	Finds documents that contain
price	The word "price"; Searching for **a single term**
price controls	The phrase "price controls"; Searching for **a phrase**
price AND discount	Both "price" and "discount"; Searching with **AND** ⬅ all caps
price OR discount	Either "price" or "discount"; Searching with **OR** ⬅ all caps
price AND NOT discount	"Price" but not "discount"; Searching with **AND NOT** ⬅ all caps
(total profit) AND (revenue OR income)	The phrase "total profit" and either "revenue" or "income"; **Combining Boolean operators**(AND, NOT, OR)
"profit and loss"	The phrase "profit and loss" (**double-quotes** required for phrases that contain search operators)
succe*	Any word that begins with "succe," such as "success," "succeed," and "successful"; Searching with **wildcard characters**
report??	Any eight-character term that begins with "report," such as "report08," "report89," and "reporter."

TOP

Using Search Options

The search options expand and restrict the results of searches with single terms, phrases, and Boolean expressions. The search options do not apply to Document Information field values.



<u>Definition of the Search Options</u>

Word Stemming: finds words that share a stem with the search term.

Thesaurus: finds words that have the same meaning as the search term.

Sounds Like: finds words that share a phonetic resemblance to a search term.

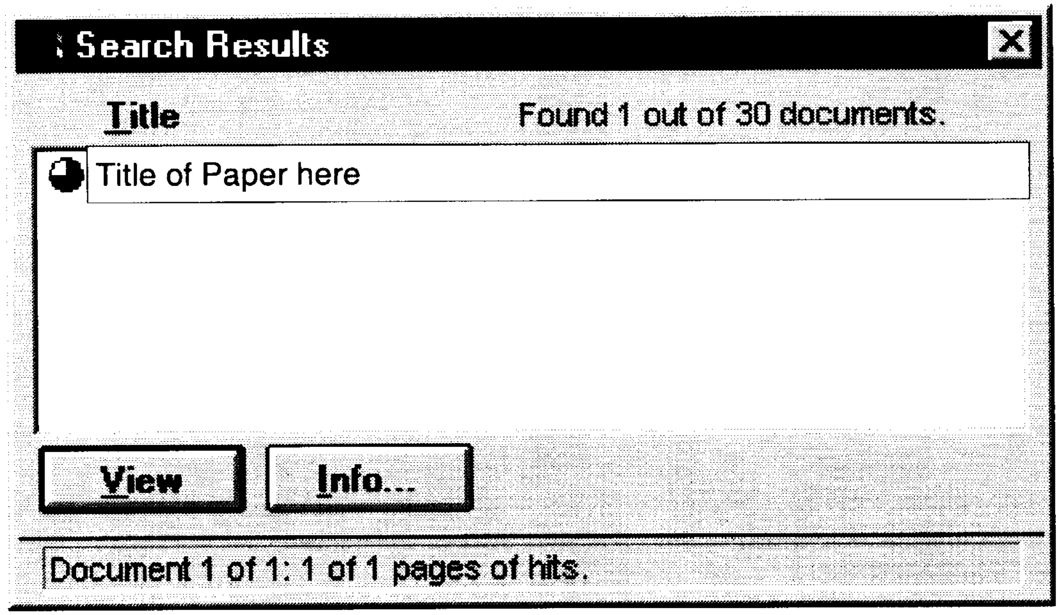
Match Case: finds only words that are capitalized exactly like the search term.

Proximity: changes AND searches so that words must be within 3 or 4 pages of one another.

TOP

Viewing Search Results

Acrobat™ Search lists the documents it finds in the Search Results window. The results are in ranking of how likely it is that the document contains the information for which you were searching (see example on next page).



Info... will give you general document information of the highlighted title.

View will open the highlighted document.
Double-clicking on the title will directly open that document.

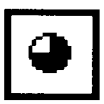
TOP

Relevance Ranking (Score)

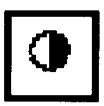
Acrobat™ Search assigns a relevance ranking to every document returned from a search. A document's relevance ranking indicates how likely it is that the document contains the information for which you are searching.

There are five icons to indicate a document's relevance ranking:

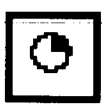
A full circle indicates a very high relevance ranking; the document is very likely to contain relevant information.



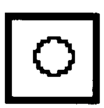
A three-quarter circle indicates a high relevance ranking; the document is likely to contain relevant information.



A half circle indicates a medium relevance ranking; the document probably contains relevant information.



A one-quarter circle indicates a low relevance ranking; the document might contain relevant information.



An empty circle indicates a very low relevance ranking; the document is unlikely to contain relevant information.

TOP

How Relevance Rankings are Determined

The method used to determine a document's relevance ranking depends on the kind of search performed.

With a **single search term**, for example, the relevance ranking indicates how frequently the term appears in the document.

When **two search terms** are separated by an OR operator, however, documents that contain both terms have a higher relevance ranking than documents that contain just one of the terms. Similarly, when you use the Proximity search option, the closer search terms are within a document, the higher the relevance ranking for the document.

Another factor that contributes to a document's relevance ranking is the relative density of search terms to other terms in the document. The higher proportion of search terms in a document, the higher the document is ranked.

Viewing the Word / Phrase from the Search Results

When you view one of the papers the search has resulted in, the word or phrase is highlighted within a specific paper. Using the scrolling search keys, you can move back and forth to the next highlighted location of the word.

1. Discuss new develop
World Wide Web, ne
[redacted] technology
meetings industry.

Actual search result of the phrase "CD-ROM"

When scrolling from highlighted word to highlighted word, the scrolling will move page to page within that particular paper .
When there are no more highlighted words within that paper, the "search next" button will take you to the next paper that contains the highlighted word/phrase. This allows you to view every highlighted word/phrase within the entire indexed group of papers.

Search Next displays the next page on which a search term occurs, or it highlights the next occurrence of a search term whether on the same or next page or the same or next paper.

Search Previous displays the previous page with an occurrence of a search term or it highlights the previous occurrence of a search term whether on the same or previous page or the same or previous paper.■

TOP

Tools Bar Icons

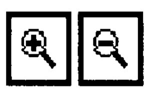
The tool bar contains tools for selecting and viewing documents. Select a tool by clicking the tool icon. To hide or show the tool bar, choose Hide Tool Bar/Show Tool Bar from the Window menu.

 Click the **Page Only** tool icon to close the overview area of the window.

 Click the **Bookmarks and Page** tool icon to open the overview area and display bookmarks created for the document. Click a bookmark's name to go to the location marked by that bookmark.

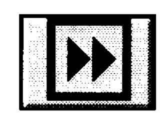
 Click the **Thumbnails and Page** tool to open the overview area and display thumbnail images of each document page. Click a thumbnail to go to the thumbnailed page.

 Use the **Hand** tool icon to move a document page on-screen when it does not fit within the window. Drag the hand tool in the direction you want to move the page.

 Use the **Zoom** tools to magnify and reduce the page display by a factor of 2.

 Use the **Select Text** tool to select text in a document, which can then be copied to the Clipboard using the Copy command. Choose **Select Graphics** from the Tools menu to select graphics in a document.

Main Menu



The **Browse** tool icons advance the document to the first or last page of a document, or forward or back one page at a time.

Use the **Go Back** and **Go Forward** tool icons to retrace your steps in a document, moving to each view in the order visited. Go Back also returns you to the original document after you click a link to another document.

Click the **Actual Size** tool icon to display the page at 100 percent.

Click the **Fit Page** tool icon to scale the page to fit within the window.

Click the **Fit Width** tool icon to scale the page width to fill the width of the window. Pressing this button and the Option (Macintosh) or Control (Windows) key fills the window with only the visible text and graphics on the page.

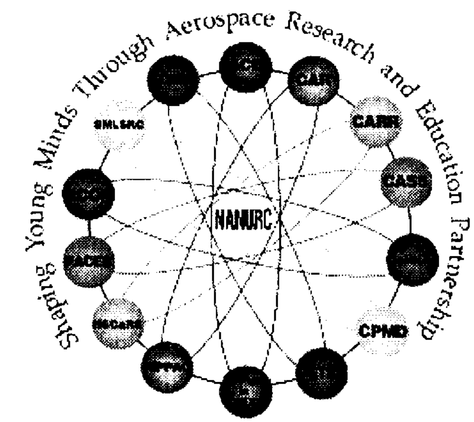
The **find** tool (binoculars only) searches for part of a word, a complete word, or multiple words in a document.

The **search** tool (binoculars with paper behind it) searches indexed PDF documents for single words, phrases, parts of words, orcombinations of words.

The **search/results** tool displays the results of your cross-document search.■

TOP

NASA University Research Centers



Alabama A&M University, Normal, AL
Center for Hydrology, Soil Climatology and Remote Sensing - HSCaRS
Morehouse School of Medicine, Atlanta, GA
Center for Space Medicine and Life Sciences - SMLSRC
Prairie View A&M University, Prairie View, TX
Center for Applied Radiation Research - CARR
Tennessee State University, Nashville,TN
Center for Automated Space Science - CASS
The University of New Mexico, Albuquerque, NM
Center for Autonomous Control Engineering - ACE
The University of Puerto Rico, Mayaguez Campus
Tropical Center for Earth and Space Studies - TCESS
The University of Texas at El Paso
The Pan American Center for Earth and Environmental Studies - PACES
Clark Atlanta University, Atlanta, GA
High Performance Polymers and Ceramics Research Center - HIPPAC
Fisk University, Nashville, TN
Center for Photonic Materials and Devices - CPMD
Florida A&M University, Tallahassee, FL
Center for Nonlinear and Non equilibrium Aeroscience - CeNNAS
Hampton University, Hampton, VA
Research Center for Optical Physics - RCOP
Howard University, Washington, DC
Center for the Study of Terrestrial and Extraterrestrial Atmospheres - CSTEA
North Carolina A&T State University, Greensboro, NC
Center for Aerospace Research - CAR
Tuskegee University, Tuskegee, AL
Center for Food and Environmental Systems for Human Exploration of Space - CFESH

Main Menu

MAIN MENU

TABLE OF CONTENTS

PAGE 1 OF INDEX

PAGE 1 OF INDEX

PAGE 1 OF INDEX

PAGE 1 OF INDEX

NASA

PAGE 1 OF INDEX

PAGE 1 OF INDEX

PAGE 1 OF INDEX

PAGE 1 OF INDEX

PAGE 1 OF INDEX

PAGE 1 OF INDEX

PAGE 1 OF INDEX

PAGE 1 OF INDEX

PAGE 1 OF INDEX

URC97001

S/ /

Jerk Minimization Method for Vibration Control in Buildings

Ayo O. Abatan,[1] and Leummim Yao[2]

URC-TC' 97
NASA URC Technical Conference on Education. Aeronautics. Space, Autonomy, Earth, and Environment

Abstract

In many vibration minimization control problems for high rise buildings subject to strong earthquake loads, the emphasis has been on a combination of minimizing the displacement, the velocity and the acceleration of the motion of the building. In most cases, the accelerations that are involved are not necessarily large but the change in them (jerk) are abrupt. These changes in magnitude or direction are responsible for most building damage and also create discomfort like motion sickness for inhabitants of these structures because of the element of surprise. We propose a method of minimizing also the jerk which is the sudden change in acceleration or the derivative of the acceleration using classical linear quadratic optimal controls. This was done through the introduction of a quadratic performance index involving the cost due to the jerk; a special change of variable; and using the jerk as a control variable. The values of the optimal control are obtained using the Riccati equation.

1. Introduction

Classical optimal controls in buildings typically consider the control of displacements and velocities [Loh and Ma, 1994, Soong and Yang, 1988]. There is minimum consideration for the control of acceleration [Yang and Li, 1991] or the rate of change of acceleration, otherwise known as jerk [Finney and Thomas, 1994],. This imposes limitation on the effective control of motion sickness resulting from jerks. There is therefore a need for more precision about the modeling of control forces in buildings since seismic excitation is jerky. For a building structure subjected to earthquake excitations, a jerk minimization method for vibration control is proposed.

2. Equations of motion

We consider a MDOF system of the form

$$\mathbf{M\ddot{x}} + \mathbf{C\dot{x}} + \mathbf{Kx} = \mathbf{Du} + \mathbf{Ef} \tag{1}$$

where x is the n-dimensional displacement vector, **M, C,** and **K** are respectively the n x n mass, damping and stiffness matrices. **u** is the m-dimensional control vector forces and **f** is

[1] Department of Engineering, Clark Atlanta Univ., Atlanta, GA 30314
[2] Department of Mathematical Sciences, Clark Atlanta Univ., Atlanta, GA 30314

the r-dimensional external excitation vector forces, whereas, **D** and **E** are respectively the n x m and n x r location matrices of the controls and external forces.

3. Control algorithm

We seek to minimize the performance index of the form

$$\mathbf{J} = \int_0^T (\mathbf{x}^T\mathbf{Q}_1\mathbf{x} + \dot{\mathbf{x}}^T\mathbf{Q}_2\dot{\mathbf{x}} + \ddot{\mathbf{x}}^T\mathbf{Q}_3\ddot{\mathbf{x}} + \dddot{\mathbf{x}}^T\mathbf{Q}_4\dddot{\mathbf{x}} + \mathbf{u}^T\mathbf{R}\mathbf{u})\mathbf{dt} \tag{2}$$

which is the analogous of the usual linear quadratic cost, where $\mathbf{Q}_i$, $\mathbf{i} = 1,...,3$, are symmetric positive semi-definite matrices and $\mathbf{Q}_4$ and R are symmetric positive definite matrices.

Introducing the change of variable

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \\ \ddot{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix} \tag{3}$$

and using the jerk as a control variable via

$\tilde{\mathbf{u}} = \begin{bmatrix} \mathbf{u} \\ \mathbf{w} \end{bmatrix}$, with x = w , system (1) can be transformed into

$$\dot{\mathbf{z}} = \mathbf{Az} + \mathbf{B}\tilde{\mathbf{u}} + \mathbf{Hf}, \qquad \mathbf{z}@) = \mathbf{z}_0 \tag{4}$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & \mathbf{I} & \mathbf{0} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} & \mathbf{O} \\ 0 & 0 & 0 \end{bmatrix} \tag{5}$$

is the $3n$ x $3n$ system matrix,

$$\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{MM}^{-1}\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \text{ and } \mathbf{H} = \begin{bmatrix} \mathbf{0} \\ \mathbf{M}^{-1}\mathbf{E} \\ 0 \end{bmatrix} \tag{6}$$

are $3n$ x $(n+m)$ and $3n$ x r location matrices respectively for the controls and external forces. Here O and **I** denote respectively the zero matrices and the identity matrix of appropriate dimensions. The corresponding performance index is given by

$$\mathbf{J} = \int_0^T (\mathbf{z}^T\mathbf{Q}\mathbf{z} + \tilde{\mathbf{u}}^T\tilde{\mathbf{R}}\tilde{\mathbf{u}})dt \tag{7}$$

where

$\mathbf{Q} = \begin{bmatrix} \hat{\mathbf{Q}}_1 & \hat{\mathbf{0}} & \hat{\mathbf{0}} \\ \mathbf{0} & \mathbf{Q}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_3 \end{bmatrix}$ is a symmetric positive semi-definite matrix and $\tilde{\mathbf{R}} = \begin{bmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_4 \end{bmatrix}$ is a symmetric positive definite matrix.

The usual necessary conditions for optimal control are given by

$$\frac{\partial \mathbf{L}}{\partial \tilde{\mathbf{u}}} = \mathbf{0}$$
$$\frac{\partial \mathbf{L}}{\partial \mathbf{z}} = \dot{\lambda}^T \tag{8}$$

with

$$\lambda^T(\mathbf{T}) = \mathbf{O} \tag{9}$$

and Hamiltonian

$$\mathbf{L}(\mathbf{z}, \tilde{\mathbf{u}}, \mathbf{t}) = \mathbf{z}^T\mathbf{Q}\mathbf{z} + \tilde{\mathbf{u}}^T\tilde{\mathbf{R}}\tilde{\mathbf{u}} + \lambda^T(\mathbf{A}\mathbf{z} + \mathbf{B}\tilde{\mathbf{u}} + \mathbf{H}\mathbf{f}) \tag{1o}$$

The above system yields

$$\tilde{\mathbf{u}} = -\frac{1}{2}\tilde{\mathbf{R}}^{-1}\mathbf{B}^T\lambda \tag{11}$$
$$\dot{\lambda} = -\mathbf{A}^T\lambda - 2\mathbf{Q}\mathbf{z}, \qquad AT\,(\mathbf{T}) = \mathbf{O} \tag{12}$$

This problem can be solved by the Riccati equation approach. We assume that the control is regulated by the generalized state vector, i.e., we seek a solution of the form

$$\mathbf{a}(\mathbf{t}) \;=\; \mathbf{P}(\mathbf{t})\mathbf{z}(\mathbf{t})\,. \tag{13}$$

Here P(t) is a symmetric and differentiable matrix. When the external excitation vector is neglected, P(t) satisfies the so called Riccati differential equation

$$\dot{\mathbf{P}}(t) + \mathbf{P}(t)\mathbf{A} - \frac{1}{2}\mathbf{P}(t)\mathbf{B}\tilde{\mathbf{R}}^{-1}\mathbf{B}^T\mathbf{P}(t) + \mathbf{A}^T\mathbf{P}(t) + 2\mathbf{Q} = \mathbf{O}, \qquad \mathbf{P}(T) = \mathbf{O} \tag{14}$$

The solution is obtained through backward integration in time. Since in most structural engineering applications the matrix $\mathbf{P}(t)$ remains constant throughout and drops to zero near **T, P can** be assumed constant. It follows that **P** satisfies the algebraic Riccati equation

$$\mathbf{PA} - \frac{1}{2}\mathbf{PB}\tilde{\mathbf{R}}^{-1}\mathbf{B}^T\mathbf{P} + \mathbf{A}^T\mathbf{P} + 2\mathbf{Q} = \mathbf{o} \tag{15}$$

$$\tilde{\mathbf{u}}(t) = -\frac{1}{2}\tilde{\mathbf{R}}^{-1}\mathbf{B}^T\mathbf{P}\mathbf{z}(t) \tag{16}$$

If the solution of the Riccati equation obtained is written the form

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} & \mathbf{P}_{13} \\ \mathbf{P}_{21} & \mathbf{P}_{22} & \mathbf{P}_{23} \\ \mathbf{P}_{31} & \mathbf{P}_{32} & \mathbf{P}_{33} \end{bmatrix} \tag{17}$$

where each $\mathbf{P}_{ij}$ is an n x n matrix, then

$$\mathbf{u}(t) = -\frac{1}{2}\mathbf{R}^{-1}\mathbf{D}^T\mathbf{M}^{-1^T}\mathbf{P}_2\mathbf{z}(t) = -\frac{1}{2}\mathbf{R}^{-1}\mathbf{D}^T\mathbf{M}^{-1^T}(\mathbf{P}_{21}\mathbf{x}_1(t) + \mathbf{P}_{22}\mathbf{x}_2(t) + \mathbf{P}_{23}\mathbf{x}_3(t)) \tag{18}$$

and

$$\mathbf{w}(t) = -\frac{1}{2}\mathbf{Q}_4^{-1}\mathbf{P}_3\mathbf{z}(t) = -\frac{1}{2}\mathbf{Q}_4^{-1}(\mathbf{P}_{31}\mathbf{x}_1(t) + \mathbf{P}_{32}\mathbf{x}_2(t) + \mathbf{P}_{33}\mathbf{x}_3(t)) \tag{19}$$

hence from (3), the expressions (18) and (19) can be rewritten as

$$\mathbf{u}(t) = -\frac{1}{2}\mathbf{R}^{-1}\mathbf{D}^T\mathbf{M}^{-1^T}(\mathbf{P}_{21}\mathbf{x}(t) + \mathbf{P}_{22}\dot{\mathbf{x}}(t) + \mathbf{P}_{23}\ddot{\mathbf{x}}(t)) \tag{20}$$

and likewise

$$\mathbf{w}(t) = \mathbf{x}(t) = -\frac{1}{2}\mathbf{Q}_4^{-1}(\mathbf{P}_{31}\mathbf{x}(t) + \mathbf{P}_{32}\dot{\mathbf{x}}(t) + \mathbf{P}_{33}\ddot{\mathbf{x}}(t)). \tag{21}$$

Once the matrix P is obtained and the case the generalized state vector is available for measurement the control can be designed by appropriately choosing the weighting matrices Q and **R**

4. Applications

The analysis proposed will lead to the design of more efficient controls for jerky excitation like earthquake loads and the design of better controllers to prevent motion sickness.

5. Conclusion

The concept of jerk minimization was introduced in this paper with a detail of how this promising new concept can be applied to linear systems via quadratic performance index involving the cost due to the jerk. This was done through a special change of variable and using the jerk as a control variable. The values of the optimal control are obtained using the so called of Riccati equation. This result is a fruitful step toward designing efficient controllers for the minimization of vibration in structures.

Bibliography

1. Finney, R., and Thomas, G. (1994), ***Calculus, second edition,*** Addison Wesley, Reading, MA
2. Kubo, T. and Furuta, E.(1994), "Seismic response and its stability of an ADM controlled building considering control signal delay and control force saturation," ***Proceedings of First World Conference on Structural Control,*** 1 WC SC, Vol. 3, FP4, 12-21.
3. Lin, C-F (1994), ***Advanced control systems design,*** Prentice Hall Series in Advanced Navigation, Guidance, and Control, and Their Applications, Englewood Cliffs, N.J.
4. Lob, C-H and Chao, C-H (1996), "Effectiveness of active tuned mass damper and seismic isolation on vibration control of multi-storey building," Journal of Sound and Vibration (1996) 193(4), 773-792.
5. Lob, C-H and Ma, M.J (1994), "Active -damping or active-Stiffness control for seismic excited buildings," ***Proceedings of First World Conference on Structural Control,*** 1WCSC, Vol. 2, TA2, 11-20.
6. Soong, T. T., (1990), ***Active structural control: theory and practice,*** Longman Scientific and Technical, New York, NY
7. Soong, T.T. and Yang, J.N.(1988), "Recent advances in active control of civil engineering structures," ***Probabilistic Engineering Mechanics, 3(4),*** 179-188.
8. Yang, J.N., Li, Z., and Vongchavalitkul, S. (1994), "Generalization of Optimal Control Theory: linear and Nonlinear Control," ***Journal of Engineering Mechanics,*** ASCE, 120(2), 267-283.
9. Yang, J.N. and Li, Z. (1991), "Instantaneous optimal control with acceleration and velocity feedback," ***Probabilistic Engineering Mechanics,*** Part ***2, 6(3 & 4), 204-211.***

Page intentionally left blank

ASICs Approach for the Implementation of a Symmetric Triangular Fuzzy Coprocessor and its Application to Adaptive Filtering

Saleh abdel-hafeez, Scott Starks, and Bryan Usevitch
Department of Electrical and Computer Engineering
University of Texas at El Paso

ABSTRACT

This paper discusses the implementation of a fuzzy logic system using an ASICs design approach. The approach is based upon combining the inherent advantages of symmetric triangular membership functions and fuzzy singleton sets to obtain a novel structure for fuzzy logic system application development. The resulting structure utilizes a fuzzy static RAM to store the rule-base and the end-points of the triangular membership functions. This provides advantages over other approaches in which all sampled values of membership functions for all universes must be stored. The fuzzy coprocessor structure implements the fuzzification and defuzzification processes through a two-stage parallel pipeline architecture which is capable of executing complex fuzzy computations in less than 0.55μs with an accuracy of more than 95%, thus making it suitable for a wide range of applications. Using the approach presented in this paper, a fuzzy logic rule-base can be directly downloaded via a host processor to an on-chip rule-base memory with a size of 64 words. The fuzzy coprocessor's design supports up to 49 rules for seven fuzzy membership functions associated with each of the chip's two input variables. This feature allows designers to create fuzzy logic systems without the need for additional on-board memory. Finall y, the paper reports on simulation studies that were conducted for several adaptive filter applications using the least mean squared adaptive algorithm for adjusting the knowledge rule-base.

I. Introduction

Fuzzy logic systems (FLS) have been successfully applied to a wide variety of practical problems. Notable applications have centered on areas such as control, expert systems, digital signal and image processing, and robotics [1-3]. The desire to use fuzzy logic in real-time has led to the development special-purpose fuzzy hardware systems [4]-[6]. Many of these systems require the use of high-cost VLSI fuzzy logic circuits and memory chips. Often the speed of these systems is slow due to the time it takes to retrieve and save truth values. Computational accuracy can be a drawback as well. It is established in [7] that the design of an FLS can be made easier by simplifying the internal parameters of the system. Despite these simplifications, the resulting design is still capable of supporting a wide class of applications.

The aim of this paper is to present the ASICs hardware development of a fuzzy coprocessor based upon the concept of the reduced symmetric fuzzy singleton set reference [8] which helps alleviate some of the drawbacks associated with man y current fuzzy hardware systems. This fuzzy coprocessor has the following features:

(1) two singleton inputs,
(2) one crisp output,
(3) seven symmetric triangular membership functions associated with each input,
(4) fuzzy static RAM for rule-base storage, and
(5) on-chip fuzzification and defuzzification processes.

The hardware implementation can be described using VHDL code where the schematic and the detailed characteristics of the circuit are generated using an optimization compiler by Mentor Graphics.

The fuzzy coprocessor's hardware implementation requires a 64-byte Static RAM, an 8-bit sign adder/subtracter, one 8-bit sign multiplier, and an 8-bit comparator as illustrated in Figures 3-4. The design, which contains approximately 10,000 gates, has been implemented using FPGAs Alters technology and runs at a 10 MHz clock speed. Simulation results indicate that the design can run at a 25 Mhz clock speed using 1.2μ CMOSN technology. The paper also discusses the application of the proposed architecture to problems of interference noise cancellation.

II. FLS Coprocessor Design Procedures

A discussion of the operation of the FLS coprocessor was initially presented in [8]. In the present work, we will expand upon these discussions and present modifications which lead to a more cost-effective implementation. The FLS coprocessor chip provides two inputs (xl, x_2) and one output. We denote the maximum number of membership functions by the symbol K and for the present study, the maximum value for K is 7. The membership functions are constructed such that they are symmetric triangular to the center of the domain and the domain of themselves as shown in Figure 1. The end-point pair (a_{ij}^-, a_{ij}^+) completely specifies the j^{th} membership function associated with the input xi. Thus the set of all end-point pairs

$\{(a_{ij}^-, a_{ij}^+)$, i =1, 2 and j = 1,2, ...K,

completely describes the K membership functions associated with each of the 2 inputs. The structure of the membership functions restricts the absolute value of slope to be equal for all membership functions in the same universe of discourse, U_i. Also, we note from the figure that each input will be matched to exactly two membership functions in U_i. Restrictions that we place on our design enable us to store all end-points associated with our membership functions and up to 49 rules in our knowledge base in a 64-byte static RAM.

In our design, the implication and inference operations are evaluated by using product operators. This approach has been shown to yield very good results in a number of engineering applications [9]. A centroid defuzzification scheme is used to determine the output of the FLS coprocessor chip.

The design procedure for the FLS coprocessor is outlined through the following four steps:
We begin by defining K fuzzy sets associated with each universe of discourse, U_i (i= 1,2) by specifying the end-point pairs [a;, a_{ij}^+] as described above. The corresponding rules of our knowledge rule-base are denoted as $M_{F_{ij}}^L$ (L= 1,2, . . . ,m=49). We use symmetric triangular membership functions of the form

$$M_{F_{ij}}^L(x_i) = \begin{cases} 1 - \dfrac{|xi - a_{ij}|}{C_{ij}} & \text{for } |x_i| \le C_{ij} \\ 0 & \text{Otherwise} \end{cases} \quad \text{----------.(1)}$$

where C_{ij} is a normalizing constant to control the slope
i = 1, 2
j = 1, 2, ..., K=7
L= 1,2, ..., m=49.

2) Next, we construct a set of IF-THEN fuzzy rules in the following form:

$$L, = \text{IF } x_1 \text{ is } F_{1j}^1 \text{ and } x_2 \text{ is } F_{2j}^1\text{; Then } m_1 \text{ is } Q'$$

$$\cdots\cdots\cdots \quad \text{----------(2)}$$

$$L_m = \text{IF } x_1 \text{ is } F_{1j}^m \text{ and } x_2 \text{ is } F_{2j}^m\text{; Then } m_m \text{ is } Q''$$

where y_i is the consequent associated with rule L_i. Reference [8] provides more detail for this step.

3) Construct the fiber F: U → R based on the M rules of step 2 as follows:

$$F(x) = \sum_{L=1}^{m} Q^L \prod_{j=1}^{k} \prod_{h=1}^{k} (M_{F_{1j}}^L(x_1) M_{F_{2h}}^L(x_2)) \quad \text{---------(3)}$$

In this step, we form products for each of the pairs of strengths associated with each **fuzzy** set in each universe of discourse. We note that due to the structure imposed through our fuzzy membership functions, that the majority of these products will be zero and the denominator of products will be unity. Also, we note that the terms Q^L (L=1, 2, ... 49) are free parameters and the filter is nonlinear.

4) At this step, we use the following LMS algorithm to update the filter parameters $Q^{L.}$ as specified in step 2. At each time points= 1,2, . . . we perform the following adaptation:

$$Q^L(s) = Q^L(s-1) + \alpha[O_d(s) - F(x(s))]\Big(\prod_{j=1}^{k} \prod_{h=1}^{k} (M_{F_{1j}}^L(x_1) M_{F_{2h}}^L(x_2))\Big) \quad \text{----------(4)}$$

where $\alpha \leq 1$ is our learning factor and $O_d(s)$ denotes the desired output. Minimization of the LMS cost function

$$K = E\left\{(O_d(s) - F(x(s)))^2\right\}$$

ensures that the input sequence **x[s]** optimally matches the desired output sequence $O_d(s)$ at each time point s= 1,2, ... Finally, a graphical representation shown by Fig. 1 summarizes all the previous steps.

Therefore, the filter F(**x**(s)) given in Eq. 3 can match any input-output pair [x(s) ;O_d(s)] to arbitrary accuracy by properly choosing the parameters Q^{l}. However, this is the only degree of freedom we have available during the adaptation procedure because the end-points of the symmetric triangular membership functions, (a_{ij}^-, a_{ij}^+), are chosen according to a maximum input limit before the adaptation takes place.

III. ASICs Design of the FLS Coprocessor

The FLS coprocessor layout given in Fig. 3 is based on the mathematical description presented in the previous section and the concepts of Fig. 1. We obtain **8-bit** real-time operation by processing all computations in parallel with two levels of pipelines separated by a high-speed **8-bit** storage buffer. The **8-bit** high-speed signed parallel comparator, given in Fig. 4, compares the released input value with all the end-points of the symmetric triangular membership functions. There are seven of these end-points. The comparator releases the upper and lower addresses of the matched membership values as well as the end-point which is greater than or equal to the input value.

The addresses released from the comparator are stored in a 3-bit D-type flip-flop register, where they are concatenated via a 3-bit multiplexer to generate a **6-bit** address bus. This maps to a designated rule-base location stored in the **64-byte** static RAM. The matched membership function degree α_i of the given input is calculated by subtracting the input from the release end-point and multiplying the result by the appropriate positive slope. Since we have two matched membership functions for any given input value and the membership functions are normalized to one, the other degree is simply evaluated by assessing the inverse of the first evaluated degree. Four **8-bit** D-type flip-flops are used as a temporary storage during this computation.

The computation of the second stage of the pipe-line is achieved by cross multiplying the matched degree values from the given two inputs with the appropriately retrieved rule-base values. These results are then aggregated to produce the de fuzzified crisp output according to Eq. 3. In order to speed up the computations during the two stages of the pipelines, the **8-bit** signed adder/subtracter has been designed with two stages 4-bit carry look-ahead structure while the **8-bit** signed integer multiplier is designed with Wallace trees structure, which are described in [10].

The components are designed using the VHDL language and optimized by Autologic 11 (Mentor Graphics EDA design tool). Table 1 illustrates the area and the delay of the coprocessor components optimized under smallest area for the Alters technology implementation and smallest area and fastest time for implementation using the 1.2μ CMOSN technology. The 1.2μ technology provides less delay in terms of its critical path analysis and thus allows the circuit to run at 25 **Mhz** while still providing an output every 0.55 μs. Using the 1.2u CMOSN technology, we can fabricate the circuit on a single chip with a dimension of 3*3 mm

IV. Application to Adaptive Noise Cancellation

Although the hardware of the FLS coprocessor chip design is simple, the structure itself can incorporates a wide class of applications based upon LMS adaptive filter approaches. We will describe how the structure can support applications related to interference canceling using the LMS approach.

As the name implies, adaptive noise cancellation is based upon subtracting noise from a received signal. Here the operation is controlled in an adaptive manner for the purpose of improving the signal-to-noise ratio. Fig. 2 shows the general model for an adaptive noise canceler which employs dual inputs and a closed loop adaptive feedback system. The two inputs to the system are derived from a pair of sensors: a primary sensor and a reference (auxiliary) sensor. The primary input supplies an information-bearing signal and a sinusoidal interference which are uncorrelated with one another. The reference input supplies

a correlated version of the sinusoidal interference. The input data is assumed to be real valued such that the primary input can be modeled as:

$B(n) = d(n) + A_0 \cos(w_0 n + \Phi n)$

where d(n) is an information-bearing signal which is characterized by an Autoregresive process $d[n] = v(n) - 0.8458 d[n-1]$ such that v(n) is a white-noise process with zero mean and variance $\sigma^2 = O$. Here, A_o is the amplitude of the sinusoidal interference, w_o is the normalized angular frequency, and Φ_0 is the phase. The reference input is given as $U(n)=A\cos(w_o n+\Phi)$ where the amplitude A and the phase Φ are different from those in the primary input but the angular frequency w_o is the same. Consequently, applying the adaptive process presented in Eq. 4 (section 2), the results are depicted in Graph 1 for different values for the learning factor.

V. Conclusions

By incorporating symmetric triangular membership functions, the coprocessor FLS chip offers a number of significant advantages. It does not require the use of division components. This is attributable to the symmetrical unity in the denominator of Eq. 3 (see [4] for proof). This in turn accelerates computations and minimizes the area needed to implement the chip. In addition, the symmetric triangular membership structure provides a simple and effective means of storing membership functions via their end-points. This enables us to compute strengths through trivial algebraic computations and allows for easy and fast memory access.

The coprocessor can store a knowledge rule-base of 49 rules and can produce a final output for the case of two input variables every 0.55 μs through two pipeline stages using a 20 Mhz internal clock. All computations are performed through an 8-bit data bus segmented to a 4-bit fixed point arithmetic decimal point. Though the chip is limited to a single class of membership functions and performs implication, inference and defuzzification in only one manner, it is versatile enough to support a wide range of applications. Its knowledge rule-base can be adaptively altered to achieve optimized results by employing a learning algorithm.

Acknowledgment

We wish to extend appreciation to NASA for its support of this work through cooperative agreement NCCW-0089.

References

[1] P. J. King and E.H. Madani, "The Application of fuzzy control systems to industrial process," Automatic, vol. 13, no. 3, pp. 235-242, 1977.

[2] S. -G. Kong and B. Kosco, "Adaptive fuzzy system for backing up a truck-and-trailer," IEEE Trans. Neural Networks, vol. 3, no. 2, pp. 211-223, March 1992.

[3] L. X. Wang and J. M. Mendel, "Fuzzy basis functions, universal approximation, and orthogonal least squares learning," IEEE Trans. on Neural Networks, vol. 3, no. 5, pp. 807-814, 1992.

[4] J. W. Fattaruso, S.S. Mahant-shetti, and J. B. Barton, "A fuzzy logic inference processor," IEEE J. Solid-state Circuits, vol. 29, No. 4, pp. 397-401, April 1994.

[5] H. Watanabe, W. D. Dettloff, and K. E. Yount, "A VLSI fuzzy logic controller with reconfigurable, cascadable architecture," IEEE J. Solid-state Circuits, vol. 25, no. 2, pp. 367-382, April 1990.

[6] T. Yamakawa, "A fuzzy inference engine in nonlinear analog mode and its application to a fuzzy logic control, " IEEE Trans. Neural Net., vol. 4, no. 3, pp. 496-521, May 1993.

[7] V. Catania, A. Piuliafito, M. Reuse, and L. Vita, "A VLSI fuzzy inference processor based on discrete analog approach," IEEE Trans. on Fuzzy Systems, vol. 2, no. 2, pp. 93-106, May 1994.

[8] S. M. Abdel-hafeez and S. Starks," A VLSI modified architecture for reduced symmetric fuzzy singleton set and its applications," ***in application and science of artificial neural nets II, S. K. Rogers and D. W. Puck: Editors,*** Proc. SPIE, vol. 2760, April 1996.

[9] J. M. Mendel," Fuzzy logic systems for engineering; A Tutorial," Proc. of the IEEE, vol. 83, no. 3, March 1995.

[10] J. P. Hayes (1994), Computer Architecture and Organization, 2nd ed., McGraw-Hill, New York.

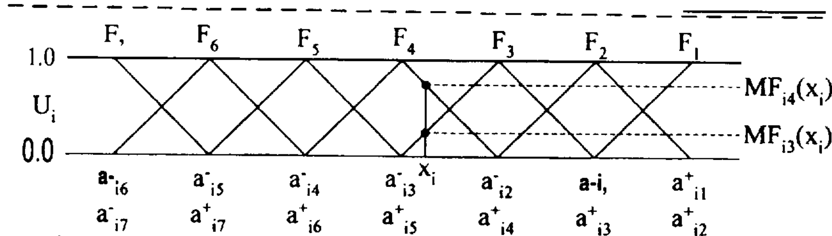


Fig. 1. Symmetric Triangular Membership Functions

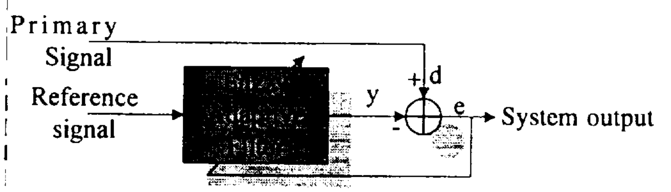
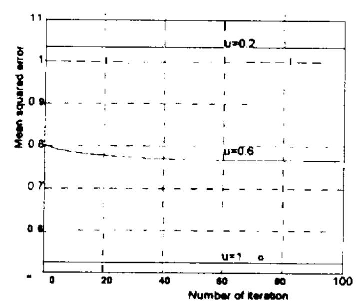


Fig. 2. A Fuzzy adaptive filtering applications for inference noise cancellation.



Graph. 1. Transient behavior of mean squared error in fuzzy adaptive fisrt-order inference noice canceller for different learning factors:

Table 1. Gate counts and maximum delav for fuzzv coprocessor in 1.2u CMOSN and Altera technologv

Coprocessor Components	CMOS Transistor Counts	ALTERA Gate Counts	CMOS Maximum Delay (ns)	Altera Maximum Delay (ns)
8-bit Signed Adder/Subtracter	310	45	15.433	28.9
8-bit Signed Special Comparator	2,104	269	19.998	40.9
8-bit Signed Multiplier	2,126	367	29.75	74.7
6X64-bvte Signed RAM	28, 230	3329	10.1	10.4

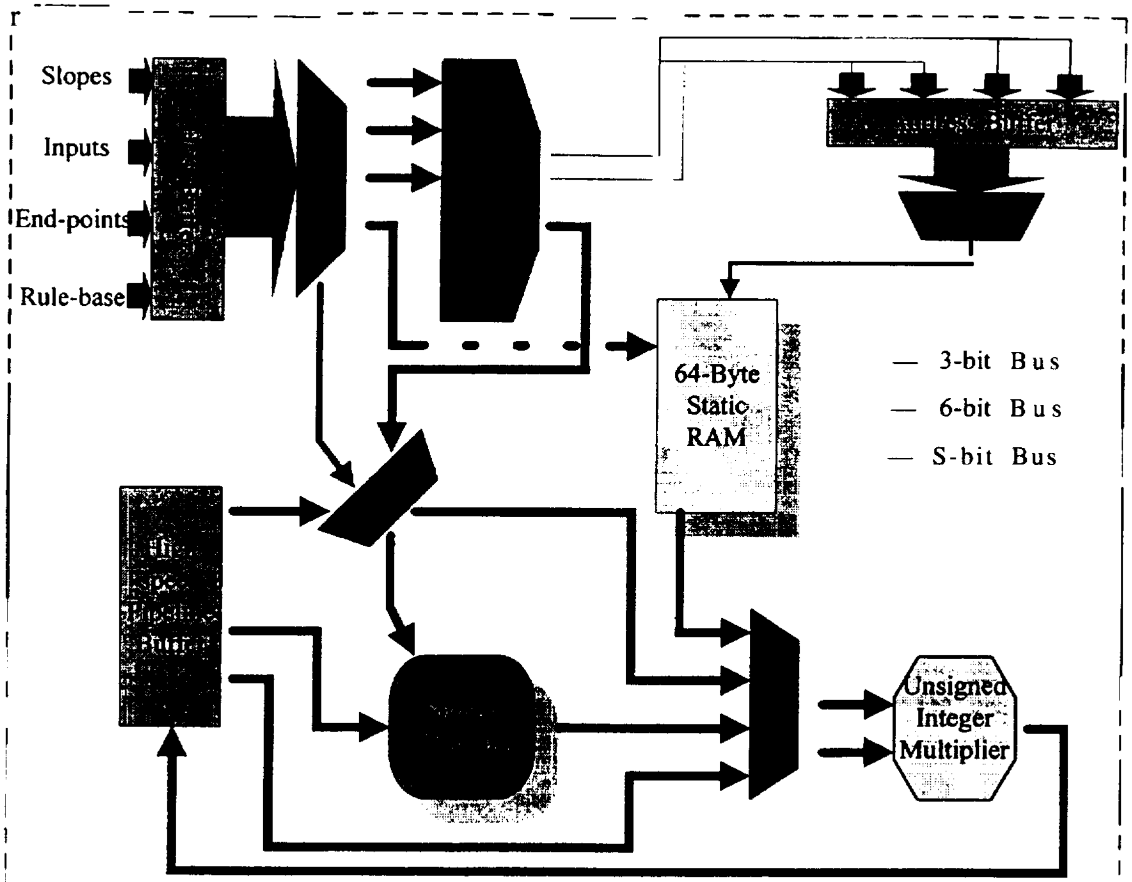


Fig. 3. Block diagram for a fuzzy coprocessor chip

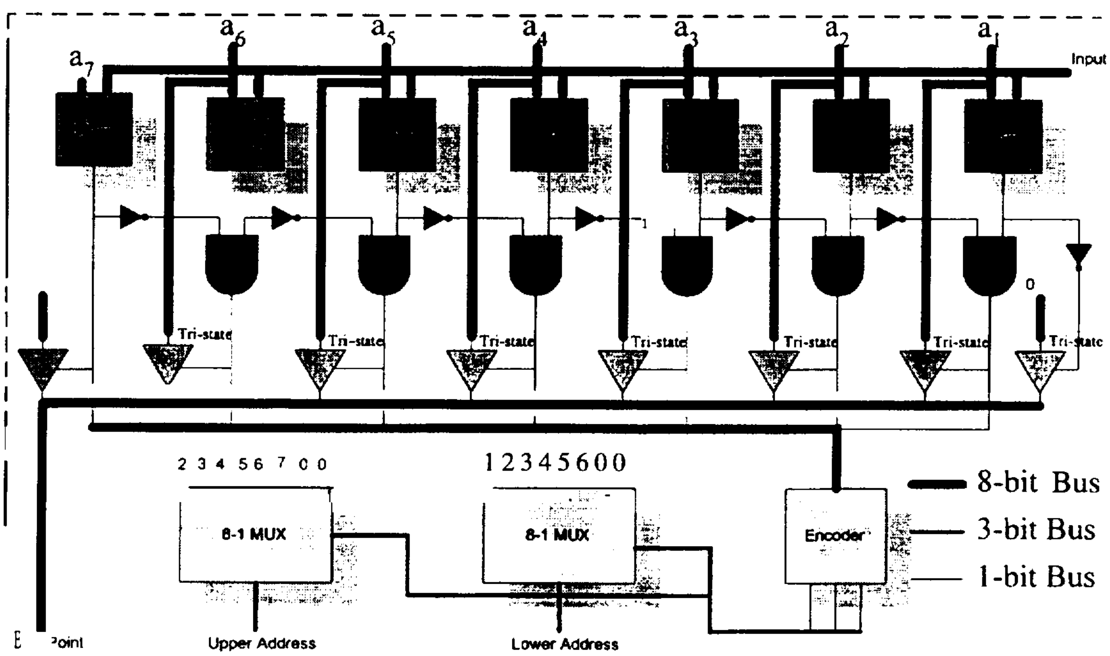


Fig. 4. Block diagram for high speed 8-bit signed comparator

URC 97003

Issues in Defining Software Architectures in a GIS Environment,

Jesus Acosta and Lori Alvarado
Department of Computer Science
The University of Texas at El Paso
El Paso, TX 79968
{jacosta,alvarado}@cs.utep.edu

1 Introduction

The primary mission of the Pan-American Center for Earth and Environmental Studies (PACES) is to advance the research areas that are relevant to NASA's Mission to Planet Earth program [1], One of the activities at PACES is the establishment of a repository for geographical, geological and environmental information that covers various regions of Mexico and the southwest region of the U.S. and that is acquired from NASA and other sources through remote sensing, ground studies or paper-based maps. The center will be providing access of this information to other government entities in the U.S. and Mexico, and research groups from universities, national laboratories and industry.

Geographical Information Systems (GIS) provide the means to manage, manipulate, analyze and display geographically referenced information that will be managed by PACES. Excellent off-the-shelf software exists for a complete GIS as well as software for storing and managing spatial databases, processing images, networking and viewing maps with layered information. This allows the user flexibility in combining systems to create a GIS or to mix these software packages with custom-built application programs. Software architectural languages provide the ability to specify the computational components and interactions among these components, an important topic in the domain of GIS because of the need to integrate numerous software packages.

'This paper discusses the characteristics that architectural languages address with respect to the issues relating to the data that must be communicated between software systems and components when systems interact. The paper presents a background on GIS in section 2. Section 3 gives an overview of software architecture and architectural languages. Section 4 suggests issues that may be of concern when defining the software architecture of a GIS. The last section discusses the future research effort and finishes with a summary,

2 Background

A Geographical Information System (GIS) is a computer-based system capable of assembling, storing, manipulating and displaying geographically referenced information, i.e., data identified according to its location. GIS technology can be used for scientific investigations, resource management, and development planning all within the scope of PACES. The types of analysis that this system must provide are measurement, mapping, monitoring, and modeling usually called the four M's of a GIS [7]. *Measurement* is concerned with storing data with various environmental parameters. A GIS must be able to retrieve old data and collect new data both efficiently and accurately.

The information may be combined with data from different sources to provide data about a region. A concern, then, is the ability to form comparisons on separate types of data. The feature of GIS that deals with integrating different types of data is referred to as *mapping*. For example, a system may allow one to relate information about the amount of rainfall to an aerial photograph of a region, providing information about which areas of the region tend to dry Up at certain times of the year. Mapping data from numerous data sets permits more intense studies of regions to be conducted.

A GIS should also allow a user to *monitor* changes in data over periods of time. For example, if a GIS can store data measurements taken at different points in time, a researcher can monitor changes that may occur in these areas.

Modeling allows a user to create models based on relations, For example, given an aerial photograph of a region and elevation readings taken from remote sensing, a model of a region could be produced allowing a more detailed study of the landscape.

Systems that provide the features discussed here can be created using off-the-shelf software. Software packages such as ESRI's ArcInfo[1], exist that provide all of the features in one package. ArcInfo is a series of six integrated software modules that provides basic GIS tools and utilities for cartographic design, query data entry and editing, data translation. polygon overlay and buffering (image processing), network analysis, and modeling. A GIS can also be constructed by combining the different, software packages that provide features necessary in such a system. A database management system such as Oracle7 can be used to store, manage and query spatial or multidimensional data. Image processing software, such as PCI's ACE[2], provides a complete environment for producing digital maps, integrating both vector and raster information. In addition, it includes a complete cartographic editing environment. Networking software such as Novell could also be used to create a GIS. An issue that is addressed in a later section is the need to specify the interface of these packages in order to provide the necessary information for successfully integrating software in this environment.

3 Software Architecture

Software architecture, an area in software engineering that has emerged seriously in the last decade, deals with the ability to formally describe the behavior of software systems. The software architecture of a system is defined in terms of computational components and the interactions among those components. When constructing a system, software architecture defines the different functionalities of the components that make up the system. This can be used to analyze the system and to determine whether the system will function properly. When integrating software systems, the software architecture is imperative when determining if the system accomplishes what is intended,

The languages that, are used to specify the architecture of software systems are called Architectural Description Languages or ADL's [4]. In order for an ADl, to be effective, three common characteristics are traditionally addressed. The first of these is the components or modules of a system. These are the parts of a system that perform actions or functions. Components are such things as clients and servers, databases, filters, and layers in a hierarchical system [5]. The second characteristic is the connections or interactions between the specified components. Interactions among components at this level of design can be things such as procedure calls and shared variable access as well as any data that is passed between components of the system. The final characteristic that is addressed in an ADL are constraints. Constraints are usually expressed on the interface of a component and are used to define restrictions.

Experimental languages for defining architectures include Rapide, MetaH and LILLEANA [8]. One of the architectural specific. at, ion language that we are examining is Rapide. A component in Rapide is specified through its interface, behavior, and constraints [3]. The interface identifies the types of events that a component can receive and generate by declaring in actions and out actions. The behavior section contains type declarations, objects, and a set of transition rules that operate on those objects and types. Objects and types model the state of the component, and transition rules model how objects and types react to actions received or generated by the component. The constraint section of a component constrains behavior by defining patterns of execution.

Consider an alarm system, as shown in Figure 1. The components of the system consist of an *Alarm Control* component, two light components and a speaker component. The light components, marked as *Ready Light* and *Armed Light*, simply display light based on input signals. The two possible input signals to these components are *off* and *on*. The speaker component is similar in function to the light components. The only difference is that the speaker component sounds an alarm instead of displaying light. The *Alarm Control* component controls the entire alarm system. The *Alarm Control* provides two kinds of external input actions: readings taken in through a sensor and codes taken in through a key pad.

An example of the AlarmControl component defined in Rapide is given below:

```
< global types >
      type signal is ..;
      on : signal := . . .;
```

[1] See http: //www.innovsys. mm/.

[2] See http://www.pci.on.ca/.

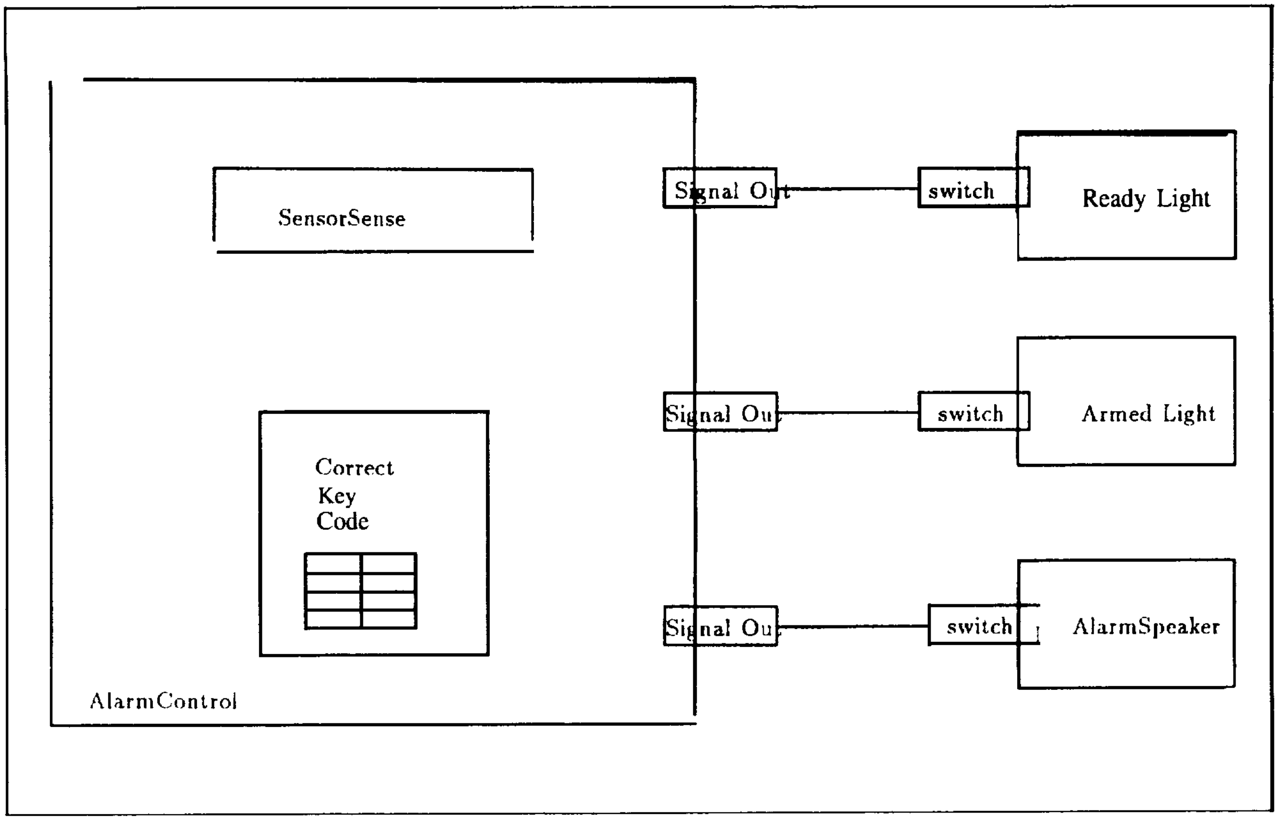


Figure 1: An alarm control component.

```
        off : signal := . . .;
interface AlarmControl is
        in action Correct_Key_Code;
        in action SensorSense;
        out action Alarm (S:signal);
        out action Armed(A:signal);
        out action Ready(R:signal);
behavior
        type state_type is (armed, ready, activated)
        state: state-type := ready;
        Correct_Key_Code where (state = ready) → state := armed;
                Armed(on) || Ready (off);;
        Correct_Key_Code where (state = armed) → state := ready;
                Armed(off) || Ready (on);;
        Correct-Key-Code where (state = activated) → state := ready;
        Alarm(off);;
        SensorSense where (state = armed) → state := activated;
        Alarm(on);;
constraint
        (
        [ Correct-Key-Code where (state = ready) →
                (Armed(on) || Ready(off)) ] <
        [ SensorSense where (state = armed) →
                (Alarm(on)) ]
        )
end Alarm Control;
```

The Rapide Language describes a component in terms of its interface, behavior within the component, and constraints on transitions that occur based on actions. [n the specification above, **in actions** arc actions that can be received from outside of the **AlarmControl** component. These include **Correct_Ke y_Code** and **SensorSense**. Actions generated with data passed out, of a component are specified as out **actions.**

Within the behavior block. actions that are generated as well as state transitions based on actions are specified. For example, if a Correct _Key_Code action is received, and the state of the system is armed, the transition from armed to ready occurs. The out actions Armed and Ready pass parameters of type **signal** out of the component. The constraint imposes an order on the actions using the < sign. The constraint above is stating that the alarm must be armed before it, can be activated.

4 Issues

There are a number of issues that are of concern when addressing systems in the GIS domain. These issues involve storage of datasets, handling of multiple formats, precision issues, and the integrity of the data in a GIS. GIS systems typically handle large amounts of data that are stored in one or more databases. When discussing software architectures in this domain, specifying how data is stored, retrieved and analyzed are central issues.

A GIS typically handles many different types of file formats, such as raster, vector, and image data. This is also a characteristic that is not, present in traditional systems. For example, if a developer wishes to use an existing off-the-shelf database to store much of the data, and an image processing software package that provides mapping features, in conjunction with a custom-built application, determining how the multiple packages store and operate on data must be communicated to the relevant stakeholders. Some of the questions that should be asked include the following [2]:

1. How are the numbers stored?
2. How is the data organized?
3. What is the dimensionality of tile data?
4. Is the data on a grid?
5. What is the best way to analyze the data?

It should be noted that numerous groups have been organized to address standards on data, for instance, compressing, holding and transmitting images. Example standards in this area include CompuServe Graphics Interchange Format (GIF) and Encapsulated PostScript Format (EPS). Standards also exist for spatial data exchange for transferring digital cartographic data anti vector/raster spatial data, e.g. the Spatial Data Transfer Specification (SDTS), US Digital Cartographic Data Standards Task Force (DCDSTF). In addition, there are standards set for database management systems [7].

An issue in using off-the-shelf packages relates to the difficulty in determining behavior of the software due to the lack of documentation; manufacturers typically do not provide documents that describe the structure of their software. Information provided in user's manuals do not address internal behavior of a system. How a package interfaces with other packages, however, can be determined through the documentation that is provided, as well as through analysis. Our focus is on defining the interfaces of packages used in GIS, and to determine if the available languages are effective architectural specification languages.

Another issue that must be considered in a GIS is the precision of data. Data used for geographical studies can vary greatly in precision. Because higher resolution of data requires considerably more storage and processing time, multiple data sets may be available for the same physical area with varying resolutions. A developer who plans to interface off-the-shelf software with a custom-built application must know how the data will vary in precision. For example, if a database package uses 32 bit numbers, and an image processing tool that will interface with the database uses 64 bit numbers, there may be a conflict which will cause loss of data when analysis occurs. Providing specifications of interfaces of the software packages that include this inform at ion would alert a developer to these types of changes.

The final issue deals with the support, of integrity checking. In a geographical information system, data integrity is critical as data is transferred from one software component to another. Invalid data input can significantly modify information used in a study. It is essential that architectures, and the languages used in the GIS arena address all these issues.

5 Future Work

The traditional view of software architectures and architectural description languages are designed to address general systems. in the domain of geographical information systems, issues exist which may or may not be addressed by the traditional view of architectures or existing architectural specification languages. The focus of the ongoing research is in further identifying design issues present in the GIS domain and determining what is necessary to describe a system architecturally. It also deals with being able to use a language to describe an off-the-shelf package in terms only of its interface.

It is not known whether this can be done effectively with existing languages. Part of the work involves examining the effectiveness of existing languages in specifying geographical information systems. Future interests include developing a survey instrument that can be distributed among software engineers and stakeholders in the GIS field. This instrument is needed to refine the issues explained above, as well as to identify additional issues that may need to be addressed in software architectures of geographical information systems.

Acknowledgments. This work was partially supported by NASA contract NCCW-0089 and NSF grant no. CDA-9522207.

References

[1] University of Texas at El Paso, "Proposal for the Pan American Center for Earth and Environmental Studies," submitted to NASA Headquarters, March 1995.

[2] Fortner, B., —-em The Data Handbook A Guide to Understanding the Organization and Visualization of Technical Data. Santa Clara, CA: TELOS, 1995.

[3] Luckham, D., et al. "Specification and Analysis of System Architecture using Rapide", IEEE Transactions on Software Engineering, 21(4):336-355, April 1995.

[4] Luckham, " Micro Rapide: An Architecture Description Language", Journal of Systems and Software, 21(3):253-265, June 1993.

[5] Shaw, M., Garlan, D., "Software Architecture", Toronto: Prentice Hall, 1996.

[6] Rapide Design Team, " Rapide 1.0 Syntax Summary", Program Analysis and Verification Group, Comput. Syst. Lab., Stanford Univ., version 1. ed., August 1993.

[7] Worboys, M., "GIS: A Computing Perspective", Bristol: Taylor and Francis, 1995.

[8] Vestal, S., " A Cursory Overview and Comparison of Four Architecture Description Languages", Honeywell Technology Center: 18 Feb. 1993.

[9] Wiederhold, G., Wegner, S., "Toward Megaprogramming", *Communications of the A CM, 35(11): 89-99, Nov. 92.*

Page intentionally left blank

SENSORY INTERACTIVE TELEOPERATOR ROBOTIC GRASPING

Keli Alark
isis@slider.unm.edu

Dr. Ron Lumia
lumia@louie2.unm.edu

NASA ACE Center for Autonomous Control Engineering (ACE)
Department of Mechanical Engineering, Room 202
The University of New Mexico
Albuquerque, New Mexico 87131

ABSTRACT

As the technological world strives for efficiency, the need for economical equipment that increases operator proficiency in minimal time is fundamental. This system links a CCD camera, a controller and a robotic arm to a computer vision system to provide an alternative method of image analysis. The machine vision system which was employed possesses software tools for acquiring and analyzing images which are received through a CCD camera. After feature extraction on the object in the image was performed, information about the object's location, orientation and distance from the robotic gripper is sent to the robot controller so that the robot can manipulate the object.

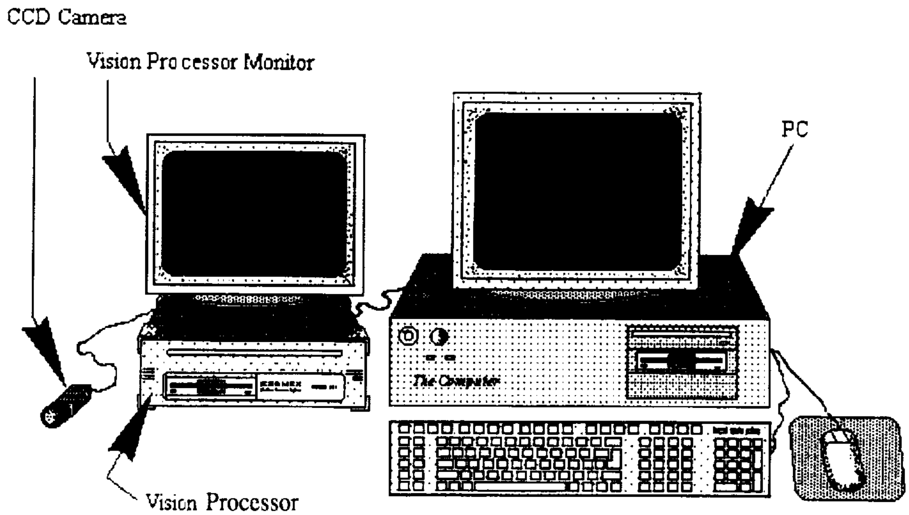
INTRODUCTION

The animal kingdom's vision and cognitive qualities has evolved to a point that sensing and reacting to stimuli in its environment are almost instantaneous [1]. As the environments of robotic systems become more complex and often dangerous, it is desired to create a system that can operate continuously in almost any environment. NASA's future goals includes the development of a permanent manned space station and the technology to insure the availability of affordable and competent satellite communications [2]. This increases the need for low-cost, low-maintenance robot vision systems that are both more mobile and more versatile. Although the processors are becoming increasingly more powerful yet more minuscule, conventional image analysis utilizes equipment that is expensive and difficult to operate. The vision system needs to be simple enough to understand in both the programming and execution phases, with minimum time expended to learn and execute each phase. The system should be compact for portability, yet the reduction in physical size of the equipment should neither increase the cost, nor decrease the capabilities of the system. However, it also needs to be versatile enough that it can be used for a multitude of different environments with minimal human operation and minimal maintenance. As a vision system is typically equipped with a camera and a CPU complete with memory and a processor adequate to analyze the data, it should be possible to locate an object, identify it by feature extraction, move to the object and manipulate it.

PROJECT-ELEMENTS

The project entailed manipulation of a pipe in an semi-random position (pipes could not be viewed on their ends) by sending the information learned from the vision system which pertained to the location, size and orientation of the pipe in real-life coordinates to the controller. Although a current trend in robotic vision is to allow camera motion in a plane or in a 3 dimensional fashion [1,3], the employed system has more thorough integration of human and computer qualities, and reduces cost. The system that was employed in this research combines a computer vision system, a robot controller, and a single CCD camera mounted to a teleoperator manipulator robot [4]. Considering that there are a myriad of vision systems to choose from, finding a system that would perform sufficiently to complete the tasks was not especially difficult. The system that was chosen was the Checkpoint 400 system by Cognex. The Cognex Checkpoint system is used for fabricating and performing vision applications such as assembly-line part inspection. The software provides the user with tools for feature extraction of the object based on the object's physical geometric properties, and its texture created from light reflection. In addition to calibration into real-world units, the software also provides operator-interface tools to facilitate the execution of the program in the workspace. Thus it was decided that although the system was not intended for precisely this type of research, it was relatively easy to program the system to performing as desired.

For the development of the programs, the system requirements include a PC with a monitor, and the CCD camera which are to be interconnected to the Checkpoint CPU Vision Processor and the Vision Processor Monitor. The PC allows the user to create and edit the text of the programs, and then debug and execute the program as necessary. After the program is completed the Checkpoint Vision Processor and the Monitor can be taken to the work-site as a stand-alone system.



There were several systems that were integrated with the Checkpoint system. Cimetrix offers an open architecture robot controller and software that allowed the application to be simulated before it was actually executed. The controller accepts inputs from the Checkpoint system in the form of data strings. Once the information from Checkpoint was sent to the controller, the Cimetrix software simulates the series of motions the robot will undergo and if the simulation proved that there were no errors, the actual robot was deployed.

Barrett Technology created a three fingered dexterous gripper. The design has one opposable finger that acts as a thumb and is un-movable. The other two fingers have a 180° range of motions. Each finger is rated for 19.62 Newtons of force that allows the Barrett Hand to grasp firmly. Its ability to grasp up to 647.46 N (especially near pivot axis of j oints) decreases the possibility that there will be injury caused to the hand or other elements of the system by a large couple-moment. The robot that is used is a Puma 560 by Unimate, which has 6 degrees of freedom.

SYSTEMATIC APPROACH TO THE ALGORITHMS

The camera functions as the eye of the system which is ultimately controlled by the operator, and as the robot moves within a workspace, the camera sends live images of the area, in real-time, back to the Vision Processor Monitor. Once the operator designates an object of interest, a still image of the object is acquired through Checkpoint which appears on the Monitor screen. The operator at the Checkpoint system then begins the algorithm which graphically extracts features of the object. The operator adjusts the Checkpoint vision tools so they highlight the features of the object, and once this is inputted, the Checkpoint computer sends data strings to the robot controller, which deploys the robot to move to the object, grasp the object, and move the object to a new location.

The Cognex Checkpoint system performs visual inspections, usually of identical or similar objects, based upon the features extracted by graphical tools pre-programmed into the Checkpoint software. The following are the steps involved in creating a repetitive recognition program. The test for the system utilized PVC pipes (due to their low density) that were either black or white:

- ♦ Define an archetype in an initial program, based on the geometric features of the object. 'This entails positioning the object in the camera's field of view, in-focus and unoccluded. The programmer accepts the acquired still-image of the object.
- \+ The programmer decides what are the best tools to accurately and consistently locate the object and calculate the necessary information,
- ♦ With each test, edge tools, which identify edges based on 1 ight and dark pixels, located opposite edges of the pipe and two lines tools were placed down either profile of the pipe. In this application, the

angle between the lines was measured to determine the orientation of the pipe, and as the diameter of the pipe was known. the diameter of the pipe on the screen would vary with distance and angle from the camera.

- ♦ Code is added to calculate the location, orientation and position of the object from the robot.

Below are the steps used each time the program is executed:

- ♦ The robot is moved in order to allow the camera to scan the area for suitable objects as defined above.
- ♦ The executable program prompts the operator to designate an object.
- ♦ Checkpoint displays the tools that were used in the program on the Vision Processor Monitor.
- ♦ The operator moves only the edge tools defined in the initial program to accommodate the new position of the object.
- ♦ The program configures the remaining tools which are dependent on the edge tools and then displays al I the chosen graphical took that were used to extract the features of the prototype on the screen.
- ♦ The data from the vision tools are calculated, displayed on the screen, and sent in data strings to the robot controller.
- ♦ Cimetrix controller deploys robot to grasp and manipulate object.

CALIBRATION

Once the camera was secured to the robot, initial measurements of the camera were taken. Both the camera and the Checkpoint system were tested to determine the amount of accuracy in the measurements. The camera acquired several pictures of objects of known sizes, at predetermined distances. The lighting of the objects was kept as controlled as possible, as shadows and specular reflections created invalid data. Next it was necessary to determine the position of the center of the field-of-view of the camera so that the object could be placed as close to the center of the field-of-view as possible to avoid spherical aberration.

Then an image of the pipe was acquired. [n an ideal environment, the pipe will have no dimples or marks on it that would prevent its profile from being rectangular. As the length of the pipe increases from the center of vision, the profile of the pipe no longer seems parallel as the lines that define the pipe's profile appear to converge. The angle to which the lines converge depends on the distance that the pipe is from the camera vision sensor, the diameter of the pipe, and the angle that the major axis of the pipe makes with the axis of the center of field-of-view. The diameter of the pipe in the image was determined at two locations across the converging profile lines, the angle of which was determined by utilizing two edge tools and the line tools in Checkpoint.

The relative position between the camera and the robot was determined by a 4 x 4 matrix and a calibration square. The calibration square measured 10 x 10 cm and a picture of the square was analyzed in Checkpoint. Each of the vertices of the calibration plate in terms of the camera were sent to the controller. The controller then moved the robot to and contacted each of the vertices and the robot's joint angles at each point was recorded. All information was then entered into a 4x4 matrix and the relative position between the camera and the robot was thus determined.

ISSUES AND RESOLUTIONS

Certain precautions had to be attended to initially. The camera had to be secured to the robot in a manner so the camera would not be hit during in the extreme ranges-of-motion of the robot gripper. As pipes can vary in length, diameter and density, the BarrettHand three-fingered gripper was chosen to be certain that the robot would have a sufficient grasp on the pipes and not damage any of the equipment by manipulation. In addition, it was necessary to ensure that the pipes were cylindrical tubes with a planer view of a rectangle and the opposite ends circles. To avoid spherical aberration, it is necessary that the image of the calibration square is as close to the center of the field-of-view as possible, without being so small as to cause errors in pixel overlapping.

The camera can be calibrated in real world coordinates utilizing the screen resolution (in pixels per inch) or by using precisely machined calibration plates at known distances. All subsequent measurements (except angle measurements) are given in units chosen by the operator. Angles are measured only in degrees (although trigonometric functions use only radians).

Calibration in general is prone to errors both mechanical and human. One method of avoiding the probable errors is by removing it completely as did Yoshimi and Allen [3]. However this there is still a calibration of the camera to the robot by using the Jacobian. Lighting was an unavoidable and immense task. The information

calculated by the system can vary greatly with disturbances in lighting. Large quantities of concentrated light can create specular reflections on even optically flat surfaces. Excessively little light will produce no image.

CONCLUSIONS

[integration of the systems above have yielded a system that is extremely versatile. First, it decreases the time an operator would invest in learning and executing the programs. Second, it limits the chances for operator error by requiring only minimal inputs from the operator and all calculations are performed within the system. Third, the system is compact and relatively inexpensive. Fourth, the system can be configured for distance learning, for institutions that cannot afford the capital investment. Finally, sensory based robotic systems could be very valuable for space exploration. For example, a Mars rover must have local self-preservation skills or wait the 17-40 seconds for communication from earth.

ACKNOWLEDGEMENTS

This work was supported in part by NASA under contract # NCCW-0087

REFERENCES

[1] J.L. Crowley, 1-1, I. Christensen, Vision as Process, Basic Research on Computer Vision Systems, Springer-Verlag, Berlin, 1991

[2] National Aeronautics and Space Administration, *Job Choices in Science and Engineering: 1997*, National Association of Colleges and Employers, pp. *94-95, 1996*

[3] B. H. Yoshimi, P.K. Allen, "Active, Uncalibrated Visual Servoing". In *Proceedings, IEE international Conference on Robotics and Automation, Volume 1*, pp. 156-161,8-13 May 1994

[4]. W. E. Snyder, Industrial Robots: Computer Interfacing and Control, Prentice Hall, Inc., Englewood Cliffs, NJ 1985

URC97005

Fixed Future and Uncertain Past: Theorems Explain Why It Is Often More Difficult To Reconstruct the Past Than to Predict the Future

Götz Alefeld[1], Misha Koshelev[2], and Günter Mayer[3]

35/61

[1]Institut fur Angewandte Mathematik
Universität Karlsruhe
D-76128 Karlsruhe, Germany
email goetz.alefeld@math.uni-karlsruhe.de

[2]Center for Theoretical Research and its
Applications in Computer Science (TRACS)
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
email mkosh@cs.utep.edu

[3]Fachbereich Mathematik
Universität Restock
D-18051 Restock, Germany
email guenter.mayer@mathematik.uni-restock.de

Abstract

At fist glance. it may seem that reconstructing the past is, in general, easier than predicting the future, because the past has already occurred and it has already left its traces, while the future is still yet to come, and so no traces of the future are available.

However, in many real life situations, including problems from geophysics and celestial mechanics, reconstructing the past is much more computationally difficult than predicting the future.

In this paper, we give an explanation of this difficulty. This explanation is given both on a formal level (as a theorem) and on the informal level (as a more intuitive explanation).

1 A paradoxical fact: in some situations, it is easier to predict the future than to reconstruct the past

At first glance, the past must be easier to reconstruct than the future. At first glance, it seems like reconstructing the past must be computationally easier than predicting the future, because:

- the past is already there, with all its traces left for the researchers to pick, while
- the future is yet to come, and it has not left any traces left.

In reality, it is often easier to predict the future. However, in many situations, it is much computationally easier to predict the future than to reconstruct the past. For example:

- In *geophysics*, if we assume that we know the laws describing how the system changes in time, then:

  predicting the *future* is reasonably easy: it means that we apply these known laws to predict the values of all physical quantities in all consequent moments of time. So, if we have enough data, we can predict what will happen in thousands and millions of years.

- However, if we want to use these same observations to reconstruct what happened in the *past*, the results of this reconstruction become much less certain and require much more computations.

- In *celestial mechanics*, if we know the current, positions, masses, and velocities of all celestial bodes, then:

  we can very accurately predict where they will be in the *future*; e.g., we can very accurately predict the future trajectory of the spaceship;

  however, it is much more difficult to reconstruct the *past* trajectory, e.g., to reconstruct where a given meteorite has come from; even when such a reconstruction is possible (as with meteorites traced to the Mars), the corresponding computations are much more complicated than the computations needed to predict the future.

How can we explain this "paradox"?

A side comment: from the common sense viewpoint, this "paradox" is not so paradoxical after all. Above, we gave "scientific" reasons why past should be easier to reconstruct. However, from the common sense viewpoint, predicting the past is difficult.

For example, the fact that the totalitarian regimes of Orwell's "1984" anti-utopia [12] could relatively easily suppress the past by destroying a few documents is a good indication that in general, reconstructing the past is a very difficult task.

Uncertainties: an informal explanation of the paradox. If we knew the exact equations, then, in principle, predicting the future and reconstructing the past would not be that different in complexity.

For example, if the equations are differential equations, then, since physical equations are usually invariant with respect to the change in time orientation (i. e., transformation $t \to -t$), both predicting the future and reconstructing the past mean, in mathematical terms, that we integrate the same system of differential equations.

In the simplified situation, when the equations describing how the future state $f = (f_1, \ldots, f_n)$ of the system is related to its past state $p = (p_1, \ldots, p_n)$ are *linear*: $f = Ap$, or

$$f_i = \sum_{j=1}^{n} a_{ij} p_j, \tag{1}$$

predicting the future means actually computing f_i from p_j, while reconstructing the past means solving the system of linear equations (1).

- For predicting the *future*, we need n multiplications and n additions to compute each of n quantities f_i that describe the future state. Totally, we need $O(n^2)$ computational steps.

- There exist algorithms that solve linear systems in $O(n^\alpha)$, where $\alpha < 2.5$, and it is conjectured that it may be possible to have $\alpha \approx 2$ (see, e.g., [5]). Thus, the computational complexity of reconstructing the *past* is almost the same as the computational complexity of predicting the future.

Since in case of exact knowledge, the tasks of predicting the future and reconstructing the past are of equal (or almost equal) computational complexity, the only reason why these tasks are in reality computationally different is because the actual knowledge is *not* precise, we have *uncertainties*.

What we are planning to do. In this paper, we will show that if we take uncertainties into consideration, then reconstructing the past is indeed much more complicated than predicting the future.

We will show it on the example of the simplest possible relationship between the past and the future: linear equation (1).

2 Motivations for the following definitions

How can we **describe uncertainty in p_j and f_i?** Enter intervals. Measurements are never 100% precise. Thus, if as the result of measuring a certain quantity, we get a certain value $\tilde{x}$, it does not necessarily mean that the actual value x of this quantity is exactly equal to $\tilde{x}$. If a car's speedometer shows 40 m.p. h., this does not mean that the speed is exactly 40.0000 m.p.h., it simply means that the speed is equal to 40 within the accuracy of this particular measuring instrument.

The manufacturer of the measuring instrument usually supplies it with the upper bound A for the measurement error $\Delta x = \tilde{x} - x$: in order words, the manufacturer guarantees that $|\Delta x| \leq$ A. (If no such estimate is given, then for any given measurement result, we can have arbitrary actual value of x and therefore, we can say nothing about the actual value. So, if we want to call something a *measurement*, some bound must be given.)

Sometimes, in addition to the upper bound for the error, we know the *probabilities* of different error values. However, in many real-life cases, we do not know these probabilities, and the upper bound A is the only information about the measurement error Ax that we have.

Since we are considering the simplest case (of a linear system) anyway, in the present paper, we will restrict ourselves to the simplest case when A is the only information.

In this case, if we have measured a quantity x and the measurement result, is $\tilde{x}$, then the only information that we have about the actual value is that this actual value cannot differ from $\tilde{x}$ by more than A, i.e., that this actual value must be within the *interval* $[\tilde{x}$ - A, $\tilde{x}$ + A$]$.

Coment. Computations that take this interval uncertainty into consideration are called *interval computations (see, e.g., [6]).*

First step towards formalization. In the problem of predicting the *future, we* measure the past values p_j and we try to reconstruct the future values f_i. Since the past values are obtained from measurements, we only know the intervals $\mathbf{p}_j = [\underline{p}_j, \overline{p}_j]$ of possible values of p_j.

Since we do not know the exact values of p_j, we cannot hope to predict the exact values of f_i; at best, we can hope to get some *intervals* $\mathbf{f}_i$ of possible values of f_i.

Similarly, when we reconstruct the *past, we* start with measuring the future values f_i. Thus, we start with the intervals $\mathbf{f}_i$, and we are interested in finding the intervals $\mathbf{p}_j$ of possible values of p_j.

We also need to describe uncertainties in a_{ij}. If we knew the coefficients a_{ij} precisely, then we would be able to complete the formalization. However, in many real-life situations, these values a_{ij} must also be determined from measurement, and are, therefore, also uncertain.

How can we describe this uncertainty? A natural way to find the values of a_{ij} is as follows:

- We prepare, very carefully. a state with the known values of parameters $p = (p_1, \ldots, p_n)$.
- Then, after a certain period of time, we measure the parameters $f_1, \ldots, f_n$ of the resulting state.

The resulting measurements have uncertainty in them, so, as a result, we have the *intervals* $\mathbf{f}_j$ of possible values of f_j. As a result, from the equation (1), we can only get *interval* estimates for the unknown values a_{ij}

Comment: this is where time symmetry is breaking. In the idealized case when measurements are absolutely precise, the problem is symmetric w .r. t. time reversal: from (1) we can go to a similar equation $p = A^{-1} f$ for an inverse matrix A^{-1}.

However, under a more realistic consideration, when we take uncertainty into consideration, the symmetry disappears. Indeed, we can carefully generate precise values in the present and trace how they evolve in the future, but the very fact that we are generating these values right now means that before the generation, these values did not exist, and therefore, their past "evolution" cannot be traced.

For example, we can very carefully place the spaceship at a given position, give it a prescribed velocity, and by measuring its trajectory, test where it goes, say, in one minute. However, it is impossible to make an experiment in which the initial position and velocity are fixed in such a way that the position in 1 minute is equal to the fixed point.

Now, we are ready for the formal definitions.

3 Definitions

Definition 1. *Let By predicting the future,* we mean *the following problem:*

GIVEN: n intervals $\mathbf{p}_j = [\underline{p}_j, \overline{p}_j]$, $1 \leq j \leq n$, *and* $n \times n$ *intervals* $\mathbf{a}_{ij} = [\underline{a}_{ij}, \overline{a}_{ij}]$, $1 \leq i, j \leq n$.

FIND: The intervals $\mathbf{f}_i = [\underline{f}_i, \overline{f}_i]$, $1 \leq i \leq n$, *of possible values of* $f_i = \sum a_{ij} p_j$ *when* $a_{ij} \in \mathbf{a}_{ij}$ *and* $p_j \in \mathbf{p}_j$.

Definition 2. Let By *reconstructing the past, we mean the* following *problem:*

GIVEN: n intervals $\mathbf{f}_i = [\underline{f}_i, \overline{f}_i]$, $1 < i \leq n$, *and* $n \times n$ *intervals* $\mathbf{a}_{ij} = [\underline{a}_{ij}, \overline{a}_{ij}]$, $1 \leq i, j \leq n$.

FIND: The intervals $\mathbf{p}_j = [\underline{p}_j, \overline{p}_j]$, $1 \leq j \leq n$, *of* possible *values of* p_j, *where* $f_i = \sum a_{ij} p_j$, $a_{ij} \in \mathbf{a}_{ij}$ *and* $f_i \in \mathbf{f}_i$.

4 Results

Known results of interval comput at ions show that predicting the past is indeed much more difficult. It is known that:

- the problem described in Definition 1 requires $O(n^2)$ computational steps, while
- the problem described in Definition 2 is, in general, computationally intractable (N P-hard) (see, e.g., [13, 7,8,9, 10]).

These results clearly prove that reconstructing the past is indeed a much more difficult problem than predicting the future.

Can we get an int uitive understanding of these results? The proofs of the above results are reasonably formal and not very intuitive. Since our goal is to solve a ***physical*** problem, we would like to have some more ***intuitive*** explanations why reconstructing the past is so more difficult.

These explanations are provided in the papers [11, 3, 2, 1, 4] that describe the geometry of the set of possible values of $p = (p_1, \ldots, p_n)$ in Definition 2. Namely, it turns out that:

- in the simplest case, this set is piece-wise ***linear*** [11];
- for symmetric matrices a_{ij}, it is piecewise *quadratic* [3, 2, 1]; and
- in the general case, it can be of ***arbitrary*** algebraic ***complexity*** [4].

On the other hand, the equations that describe the set of possible values of $f = (f_1, \ldots, f_n)$ is definition 1 is always *quadratic*.

This difference in algebraic complexity gives an intuitive explanation of why reconstructing past is a more difficult problem than predicting the future.

Acknowledgments. This work was partly supported by the NASA Pan American Center for Environmental and Earth Studies (PACES). The authors are thankful to Ann Gates, Vladik Kreinovich, and Scott Starks for their help and encouragement.

References

[1] G. Alefeld, V. Kreinovich, and G. Mayer, "Symmetric Linear Systems with Perturbed Input Data", In: G. Alefeld and J. Herzberger (eds.), *Numerical Methods* and *Error Bounds. Proceedings of the IMACS-GAMM International Symposium on Numerical Methods and Error Bounds, Oldenburg*, Germany, *July 9–12, 1995*, Akademie Verlag, Berlin, 1996, pp. 16–22.

[2] G. Alefeld, G. Mayer, and V. Kreinovich, "The shape of the symmetric solution set", In: R. B. Kearfott et al (eds.), *Applications of Interval Computations*, Kluwer, Dordrecht, 1996, pp. 61–79.

[3] G. Alefeld, V. Kreinovich, and G. Mayer, "The Shape of the Symmetric, Persymmetric, and Skew Symmetric Solution Set", *SIAM Journal on Matrix Analysis and Applications (SIMAX)* (to appear).

[4] G. Alefeld, V. Kreinovich, and G. Mayer, "The Shape of the Solution Set for Systems of Interval Linear Equations with Dependent Coefficients", *Mathematische Nachrichten* (to appear).

[5] Th. H. Cormen, Ch. L. Leiserson, and R. L. Rivest, *Introduction to algorithms*, MIT Press, Cambridge, MA, 1990.

[6] R. B. Kearfott and V. Kreinovich (eds.), *Applications of Interval Computations*, Kluwer, Dordrecht, 1996.

[7] V. Kreinovich, A. V. Lakeyev, and S. I. Noskov. 'Optimal solution of interval linear systems is intractable (NP-hard)." *Interval Computations,* 1993, *Pie.* 1, pp. *6--14*.

[8] V. Kreinovich, A. V. Lakeyev, S. I. Noskov, "Approximate linear algebra is intractable". *Linear Algebra and its Applications, 1996,* Vol. 232, No. 1, pp. 45–54.

[9] V. Kreinovich, A. Lakeyev, and J. Rohn, "Computational Complexity of Interval Algebraic Problems: Some Are Feasible And Some Are Computationally Intractable - A Survey", In: G. Alefeld, A. Frommer, and B. Lang (eds.), *Scientific Computing and Validated Numerics,* Akademie-Verlag, Berlin, 1996, pp. *293-306.*

[10] A. V. Lakeyev and V. Kreinovich, "NP-hard classes of linear algebraic systems with uncertainties", *Reliable Computing, 1997,* Vol. 3, No. 1, pp. 1--31 (to appear).

[11] W. Oettli and W. Prager, "Compatibility of Approximate Solution of Linear Equations with Given Error Bounds for Coefficients and Right-hand Sides", *Numer. Math., 1964, Vol.* 6, pp. 405-409.

[12] G. Orwell. *Nineteen eighty four,* Harcourt, Brace, & World, N. Y., 1963

[13] J. Rohn, V. Kreinovich, "Computing exact componentwise bounds on solutions of linear systems with interval data is NP-hard". *SIAM Journal on Matrix Analysis* and *Applications (SIMAX), 1995,* Vol. 16, No. 2, pp. 415-420.

Page intentionally left blank

PERFORMANCE OF LOGIC GATES UNDER RADIATION AND POST-RADIATION ENVIRONMENTS

John A. Alexander and John O. Attia
Center for Applied Radiation Research
Prairie View A&M University
Prairie View, Texas 77446

56/72

ABSTRACT

The effects of TID on the propagation delay of CMOS inverters, nand and nor gates were obtained using specially designed test chips. Radiation testing using Co^{60} indicates that the propagation delay of the logic gates increases with TID. Under post-radiation environment, there is a slight decrease in propagation delay for nor and nand gates.

1. INTRODUCTION

CMOS logic gates are the basic building blocks of space electronic systems, such as space vehicles and satellites. For low dose rates of 1 rad (Si)/s and 100 rad (Si)/s, Blaes and et al [1] has published results of propagation delay versus total dose for CMOS inverters. Blaes and co-workers have reported that the propagation delay of inverters is altered by total ionizing dose radiation. Gibbons and co-workers [2] obtained the propagation delay of inverters, two input NOR and NAND gates, It was shown that the propagation delay increases with total dose for devices with six to seven micron design rules. There are several attempts to design logic gates that are radiation tolerant [3, 4]. The objectives of this work is to obtain the effect of TID radiation on the basic CMOS logic gates with various PFET/NFET transconductance ratios. In addition, this work will explore the post-radiation effects on the logic gates.

2. TEST CHIPS FOR RADIATION TESTING

One of the performance indicators of logic gates is the propagation delay. The latter represents the logic gate speed. propagation delay is *the arithmetic* mean of the fall and rise times It is the minimum time needed to guarantee the logic state at an output.

Ring oscillators have been used to measure propagation delay[5], to characterize semiconductor process technology [6], to extract semiconductor device parameters [7]. The propagation delay, tpp, of an N-stage ring oscillator is given as:

$$t_{pp} = \frac{nT}{2N}$$

where:

n is the number of harmonics; *N* is the number of stages and *T* is the period of the oscillator

In this work, ring oscillators were used to determine the propagation delay of the logic gates. The basic elements of the ring oscillators were inverters, NOR and NAND gates. The NOR and NAND gates had their inputs tied together, 101-stage ring oscillators were built. An odd number of gates ensures that the ring oscillator circuit does not settle in a stable. In addition, to alleviate leaving the ring oscillator in a peculiar state during power up, a two-input NAND gate was made as an integral part of the ring circuit, One input of the NAND gate was used for initialization. The initialization input is held high during power up to initialize the state of the ring oscillator. The output of the N-stage ring circuit was connected to one of the inputs of the 2-input NAND gate

The transistors of the gates were designed to have the same length, but the width of the PMOS transistors Wp was made equal to a multiple number of the width of the NMOS transistor, Wn. Specifically, for the two-micron CMOS technology used to fabricate the test chips, the NMOS devices have the length Ln = 2.0E-6 m and width Wn = 6.0E-6 m. The PMOS devices have length Lp = 2.0E-6 m and width Wp = k Wn, where k=1, 2, and 3.

The Magic CAD tool [8] was used to design the circuit, To prevent latchup of the CMOS circuits, guard rings were placed around the p-wells [9]. To check the integrity of the design, the layout was extracted and simulated using PSPICE. When it was ascertained through simulation that the circuits behave correctly, the design was sent to NSF

MOSIS for fabrication. Three chips were fabricated using MOSIS two-micron scalable CMOS (p-well) technology.

3. TESTING METHODOLOGY

The IC chips were tested before and after total ionizing dose radiation. To check the integrity of the IC chips before TID testing, the input of 101 stage ring oscillator, was connected to a function generator and the input and output signals were observed on a scope. The output signal showed an inversion and a delay. In addition for each ring oscillator, the input was connected to an oscilloscope and a counter. The period of oscillation was observed The propagation delay was calculated using equation(l).

The effect of TID radiation on IC chips was obtained by using Co^{60}. The Co^{60} source provided dose rates of approximately 300 krad per hour. The MIL- STD- 883 proposed method 1019 [10] was used as a guide in the radiation testing. The IC chips were irradiated for 20 minutes and the propagation delay was obtained. The radiation and the propagation delay measurements were continued until the TID radiation of lMrad was obtained. After radiation exposure, the propagation delay of the devices were measured 24,48, 72,%, and 120 hours after the radiation testing.

4. TEST RESULTS

Figure 1 shows the propagation delay versus total dose for inverterl, inverter2, anbd inverter3. Inverterl, inverter2, and inverter3 are inverters with PFET/NFET transconductance ratio of one, two, and three, respectively. It can be seen from figure 2, that the propagation delay of CMOS inverterl (inverter with PFET/NFET transconductance ratio of one), and inverter2 (inverter with PFET/NFET transconductance ratio of two) are almost the same. Inverter3 (inverter with PFET/NFET transconductance ratio of three) has the highest propagation delay. Figure 2 shows the post-radiation response.

Figure 3 shows the propagation delay versus total dose for NAND21 (two-input NAND gate with unity respect ratio) and NAND22 (two-input NAND gate with PFET/NFET transconductance ratio of two). The propagation delay of NAND21 and NAND22 increases with TID radiation. In addition, the propagation delay of NAND22 is less than that of NAND21. Figure 4 shows the post-radiation response of the NAND gates.

The propagation delay versus TID radiation for three input NOR gates is shown in Figure 5. It should be noted that the propagation delay of three input NOR gate increases with TID radiation. In addition, it should be noted that the propagation delay of NOR3 1 (three input NOR gate with PFET/NFET transconductance ratio of one) and NOR32 (three input NOR gate with PFET/NFET transconductance ratio of two) have almost the same values. Figure 6 shows the post-radiation response of the NOR gates.

5. COMMENTS ON RESULTS

The results of the radiation testing indicates that propagation delay of logic gates increases with increases of TID radiation. The results of this work confirms the findings of previous investigators [2,3, 4]. Logic gates with PFET/NFET transconductance ratio of one or two have very similar propagation delays. Furthermore, logic gates with PFET/NFET transconductance ratio of three have higher propagation delays that gates with PFET/NFET transconductance ratio of one or two. It is therefore advisable to design logic gates with PFET/NFET transconductance ratio of two, since they have lower propagation delay and would almost have symmetrical characteristics [9].

The post-radiation results are novel. In the case of the inverter, there was hardly any change in the propagation delay within 120 hours after irradiation. However, there was reduction in the propagation delay with time for the NAND and NOR gates, The decrease in the propagation delay did not reach pre-irradiation values.

It is will known that there is negative shift in threshold voltage with increase in total dose. In addition, there is degradation of mobility and an increase in leakage current with total dose [8]. The *results* of this work indicate that negative shift in threshold voltage and mobility degradation with TID leads to an increase in the propagation delay of logic gates.

5. CONCLUSIONS

The results of the work show that propagation delay of logic gates increases with increases of TID radiation. This is in line with the findings of previous investigators. Logic gates with PFET/NFET transconductance ratio of a one or two have very similar propagation delays. Furthermore, logic gates with PFET/NFET transconductance ratio of one or

two have similar propagation delays. The propagation delay during post-irradiation was almost constant for inverters and decreased for NAND and NOR gates.

6. REFERENCES

[1] Blaes, B.R., Budder, M.G., Lin Y.S. *Propagation Delay Measurements from a* Timing *Sampler Intended for* Use *in Space, IEEE* Trans. Nuclear Science, NS-34, no, 6, Dec. 1987.

[2] Gibbons, C.F., Habing, D.H. and Flores, R.S. *A Radiation Hard Silicon Gate CMOS Cell Family. IEEE* Trans. on Nuclear science, Vol. NS-27, Dec. 1980, pp. 1712-1715.

[3] Chen, C, Liu, S. Hsiano, C. Hwu, J. *A Circuit Design for the Improvement of Radiation Hardness in CMOS Digital Circuits, IEEE* Trans. on Nuclear Science, Vol. 39, April 1992, pp. 272-277.

[4] Francis, P., Michel, C. Flandre, D and Colinge, J.P. *Radiation-Hard Design for SOI MOS Inverters, IEEE* Trans. on Nuclear science, Vol. *41,* April *1994,* pp. *402-407.*

[5] Moore, D.F., Holburn, D.M., Nockolds, C.E. and Rosolen, G.C. *"CMOS Ring Oscillators as Teaching Tools", IEEE* proceedings -G, vol 139, no. 2, pp. 199-204, April, 1992.

[6] Coll, P., Robert M, Reginer, X. And Auvergne, *Process Characterization with dynamic test structures,* Electronic *Letters, vol. 29,* no. *20, pp.* 1764-1766, 30th Sept., 1993.

[7] Kovacs, F. and Hosszu, G. *Semiconductor Device Parameter Extraction based on Reconfigurable Rng Oscillator frequency measurements,* Proc. IEEE 1991 In. Structures, Vol. 4, no. 1, pp. 221-224, March 1991.

[8] Ousterhout, J.K., Homachi, G.T., Mayo, N., Scott, W. S., and Taylor G. S., The *Magic VLSI Layout System,* IEEE Design and Test, pp. 47-58,1985.

[9] Uyemura, J.P. *Circuit Design for CMOS VLSI,* Kluwer Academic Publishers, 1993.

[10] MIL-STD-883, Proposed Method 1019.4, *Ionizing Radiation Effects (Total Dose) Test Procedure.*

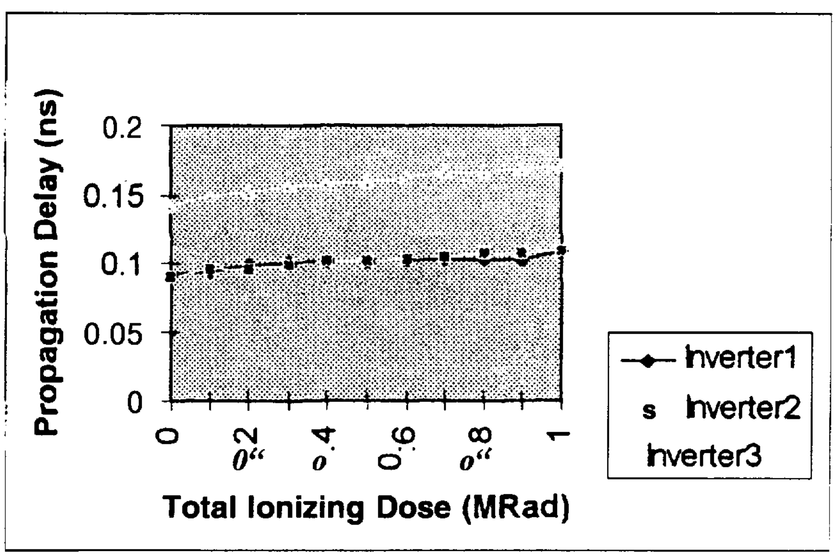


Figure 1 Propagation Delay vs. Total Dose for CMOS Inverter

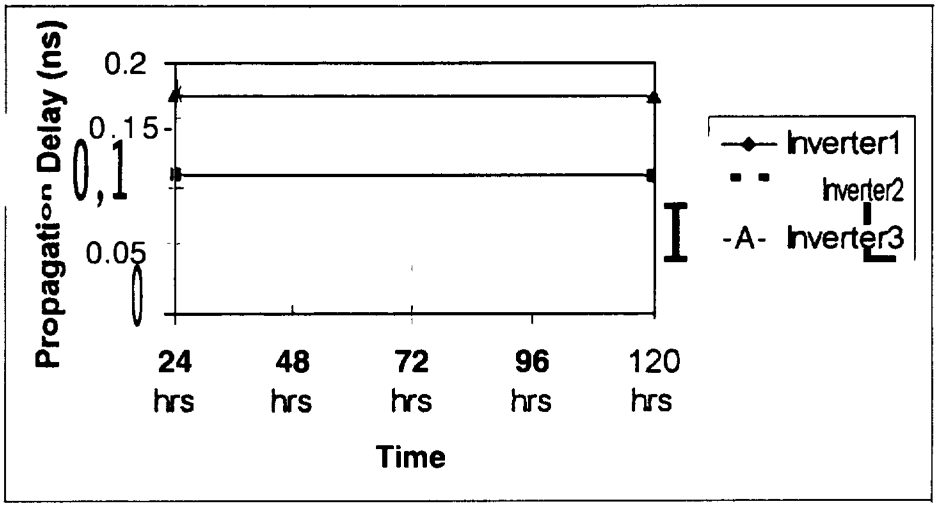


Figure 2 Propagation Delay vs. Time (Post-Irradiation) for Inverters

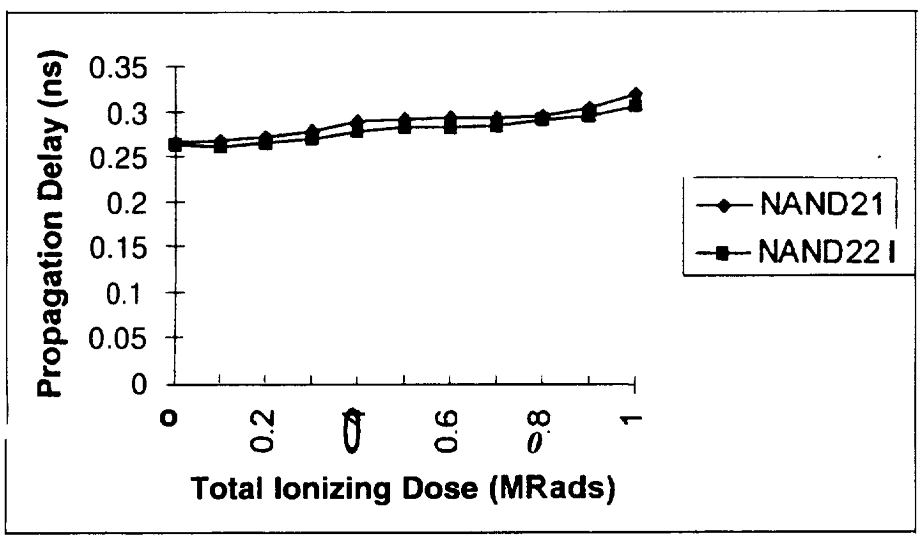


Figure 3 Propagation Delay vs. Total Dose for CMOS 2-input NAND Gate

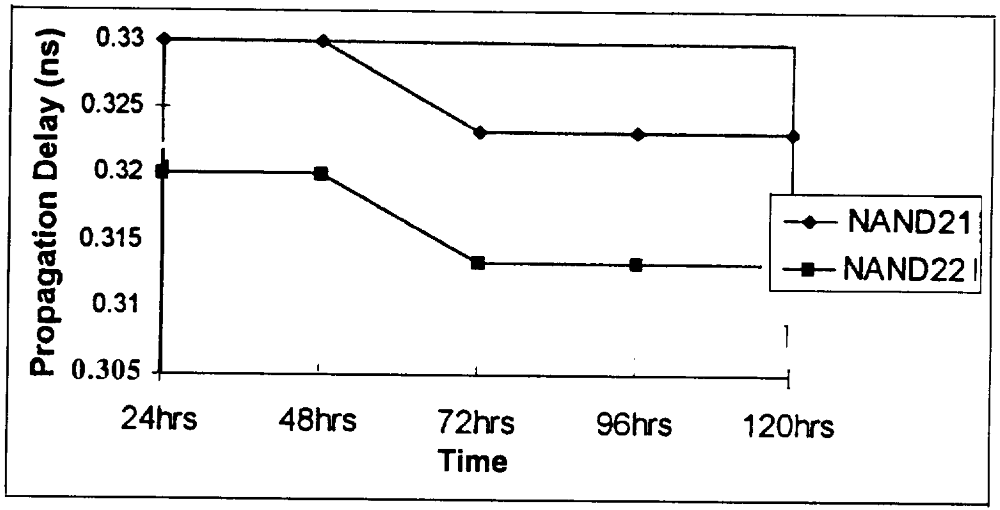


Figure 4 Propagation Delay vs. Time (Post-Irradiation) for 2-input NAND Gate

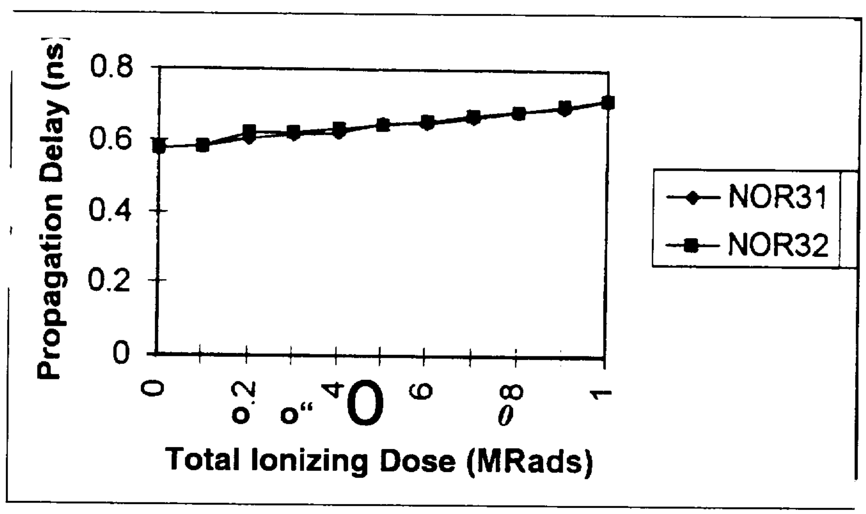


Figure 5 Propagation Delay vs. Total Dose for 3-input NOR Gate

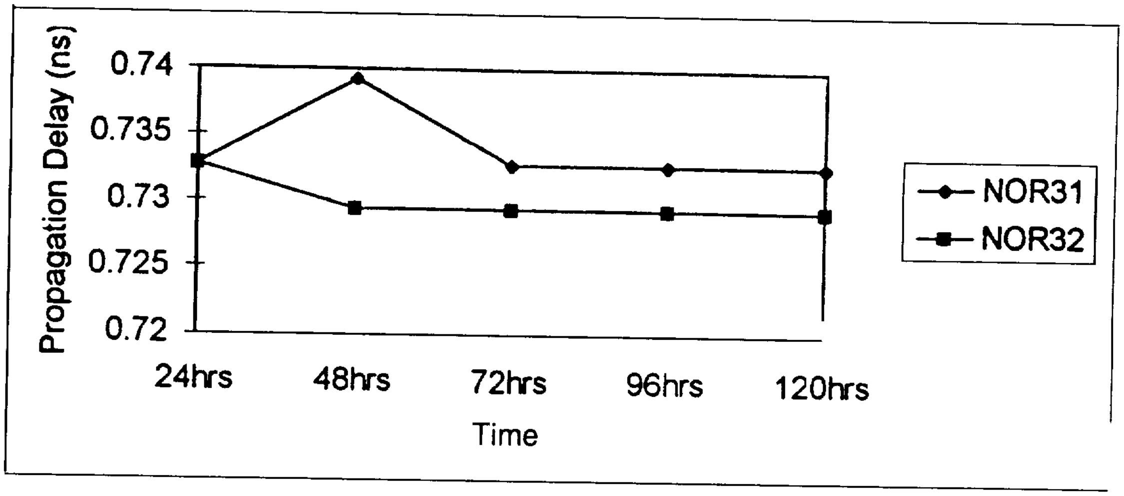


Figure 6 Propagation Delay vs. Time (Post-Irradiation) for 3-input NOR Gate

URC97007

ADVANCED TECHNOLOGY FOR ISOLATING PAYLOADS IN MICROGRAVITY

S7/21

DEAN C. ALHORN
NASA MARSHALL SPACE FLIGHT CENTER

INTRODUCTION

One presumption of scientific microgravity research is that while in space disturbances are minimized and experiments can be conducted in the absence of gravity. The problem with this assumption is that numerous disturbances actually occur in the space environment. Scientists must consider all disturbances when planning microgravity experiments. Although small disturbances, such as a human sneeze, do not cause most researchers on earth much concern, in space, these minuscule disturbances can be detrimental to the success or failure of an experiment, Therefore, a need exists to isolate experiments and provide a quiescent microgravity environment.

The objective of microgravity isolation is to quantify all possible disturbances or vibrations and then attenuate the transmission of the disturbance to the experiment. Some well-defined vibration sources are: experiment operations, pumps, fans, antenna movements, ventilation systems and robotic manipulators. In some cases, it is possible to isolate the source using simple vibration dampers, shock absorbers and other isolation devices. The problem with simple isolation systems is that not all vibration frequencies are attenuated, especially frequencies less than 0.1 Hz. Therefore, some disturbances are actually emitted into the environment. Sometimes vibration sources are not well defined, or cannot be controlled. These include thermal "creak," random acoustic vibrations, aerodynamic drag, crew activities, and other similar disturbances. On some "microgravity miss ions," such as the United States Microgravity Laboratory (USML) and the International Microgravity Laboratory (IML) missions, the goal was to create extended quiescent times and limit crew activity during these times. This might be possible for short periods, but for extended durations it is impossible due to the nature of the space environment. On the International Space Station (ISS), vehicle attitude readjustments are required to keep the vehicle in a minimum torque orientation and other experimental activities will occur continually, both inside and outside the station. Since all vibration sources cannot be controlled, the task of attenuating the disturbances is the only realistic alternative.

Several groups have independently developed technology to isolate payloads from the space environment. Since 1970, Honeywell's Satellite Systems Division has designed several payload isolation systems and vibration attenuators. From 1987 to 1992, NASA's Lewis Research Center (LeRC) performed research on isolation technology and developed a 6 degree-of-freedom (DOF) isolator and tested the system during 70 low gravity aircraft flight trajectories. Beginning in early 1995, NASA's Marshall Space Flight Center (MSFC) and McDonnell Douglas Aerospace (MDA) jointly developed the STABLE (Suppression of Transient Accelerations By Levitation Evaluation) isolation system. This 5 month accelerated effort produced the first flight of an active microgravity vibration isolation system on STS-73/USML-02 in late October 1995. The Canadian Space Agency developed the Microgravity Vibration Isolation Mount (MIM) for isolating microgravity payloads and this system began operating on the Russian Mir Space Station in May 1996. The Boeing Defense & Space Group, Missiles & Space Division developed the Active Rack Isolation System (ARIS) for isolating payloads in a standard payload rack. ARIS was tested in September 1996 during the STS-79 mission to Mir.

Although these isolation systems differ in their technological approach, the objective is to isolate payloads from disturbances. The following sections describe the technologies behind these systems and the different types of hardware used to perform isolation. The purpose of these descriptions is not to detail the inner workings of the hardware but to give the reader an idea of the technology and uses of the hardware components. Also included in the component descriptions is a paragraph detailing some of the advances in isolation technology for that particular component. The final section presents some concluding thoughts and a summary of anticipated advances in research and development for isolating microgravity experiments.

STABLE SYSTEM DESCRIPTION

The STABLE system was an experiment with a two-fold purpose: provide an isolated platform for scientific payloads and demonstrate the technology to produce isolation in a microgravity environment. STABLE isolates only the vibration-sensitive portion of a scientific payload from external disturbances. For this flight, the isolated payload was a microgravity-sensitive fluid diffraction experiment named "CHUCK." CHUCK used a combination of lasers, optics and two fluid cells, a control ceil and a test cell, to produce a diffraction pattern that was recorded on video. The diffraction pattern changed when the fluid in the test cell was heated. This pattern was very sensitive to microgravity accelerations and required isolation. To isolate the experiment, several types of components are utilized: actuators, umbilicals, sensors, control system, electronics, command input devices and a data acquisition system. To develop a 6-DOF isolation system, the sensors and actuators need to be oriented in an appropriate configuration. Figure 1 represents the configuration of the STABLE system components.

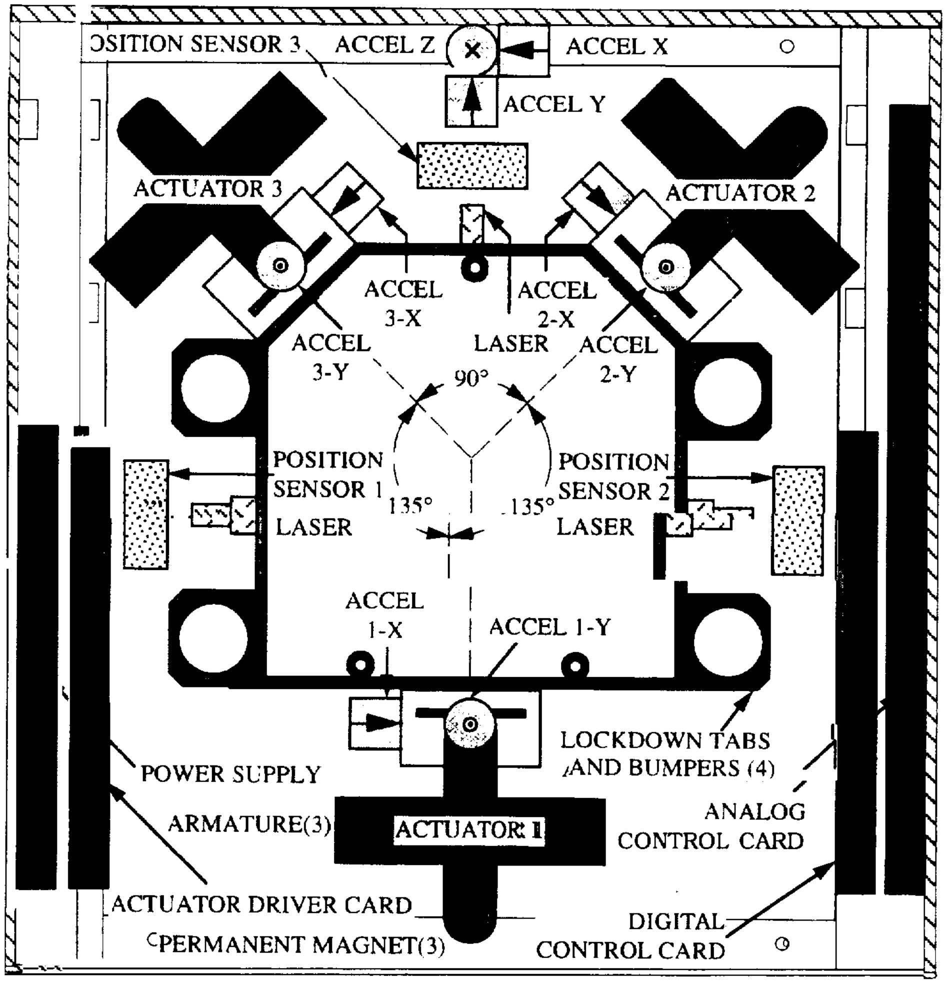


Figure 1. -STABLE System Layout Top View)

The STABLE design utilized three MDA-patented dual-axis electro-magnetic actuators. As shown in Figure 2, the actuators resemble a horseshoe-shaped permanent magnet around a paddle-shaped armature. The magnets provide a constant field which reacts to the varying magnetic field from the armature windings. For simplicity, the permanent magnets were mounted on the floated platform and the armature was mounted to the rigid base. In this mounting configuration, the control currents are not transmitted through the umbilicals. Each actuator provides two axes of force over the fill range of free travel, ±1cm. This allows the platform to "sway" in the rattlespace to achieve its microgravity performance. The forces required by the actuators, as the platform moves in the rattlespace, are directly dependent upon the umbilical design.

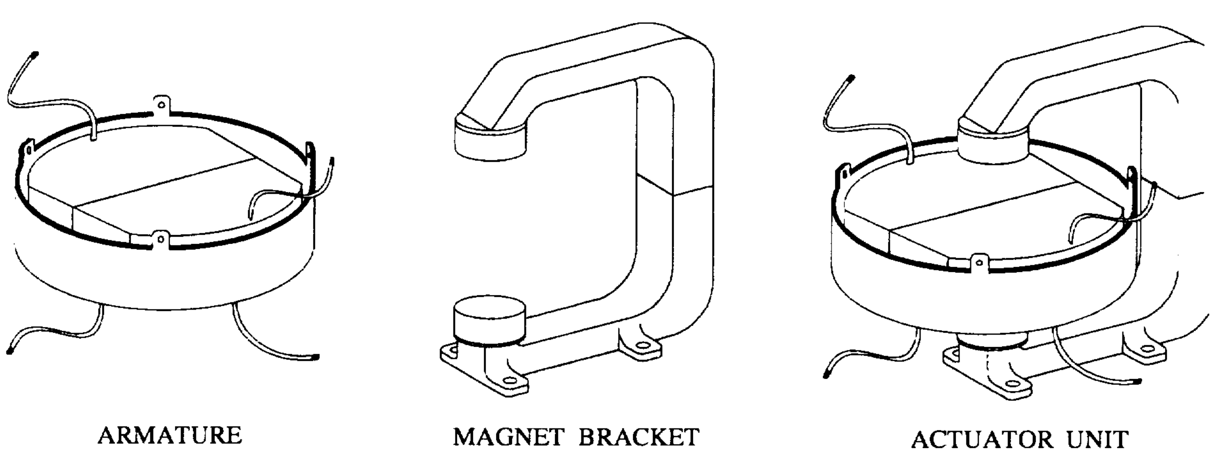


Figure 2. - MDA Dual-Axis Actuator

The STABLE umbilicals are extremely flexible cables which transmit power and signals between the base and the floated platform. Since the actuators do not rigidly connect to the platform, the only means by which the platform can be disturbed is through. the umbilicals. For this reason, the umbilicals need to be as flexible as possible. Ideally, the umbilicals should have zero stiffness, but this situation is not realistic with payloads requiring significant power, data, coolant, inert gas or vacuum needs.

For measuring position and acceleration, STABLE incorporated the SC-50D position sensor, from UDT Sensors, and the Sunstrand QA-2000, which is considered the microgravity industry standard accelerometer. These sensors were chosen for several reasons: compact size, durability, performance and relatively low cost. Of the nine single-axis accelerometers used, six were collocated with the actuators for measuring inertial platform accelerations and three accelerometers were mounted to the base to measure the nominal vibration environment for evaluating isolation performance. Three pen lasers, located on the floated platform, produced the reference light spots for the corresponding dual-axis posit ion sensors. This sensor configuration allowed STABLE to measure the relative position of the platform with respect to the base and maintain the platform in a centered position. Refer to Figure 1 for locations and orientations of the STABLE sensors.

The STABLE control system is a combination of a low frequency position controller with a high frequency analog acceleration controller with each axis having its own independent position and acceleration loops. This combination provides continuous high bandwidth, 35 Hz, acceleration control with a low bandwidth, 0.05 Hz, position control system, allowing the platform to drift relatively freely in its rattlespace. Collocating the platform accelerometers with the corresponding actuator's axis provides direct absolute feedback for the acceleration control loop. Since the position sensors are not aligned with the actuator axes, the position controller performed the appropriate transformations to determine the position errors at each actuator gap. Using this gap position information, the position controller gradually recentered the floating platform over a period of minutes while the analog control system operated continuously. The position control system normally operates in a low-gain mode. If the floated platform approached the boundaries of its sway space, the position controller would enter a high-gain mode by increasing the digital control system gains which decreases the effective acceleration control gain. This is necessary to prevent mechanical contact between the platform and base and thus maintain a quiescent microgravity environment for the sensitive payload. Figure 3 is a representation of the STABLE control system architecture.

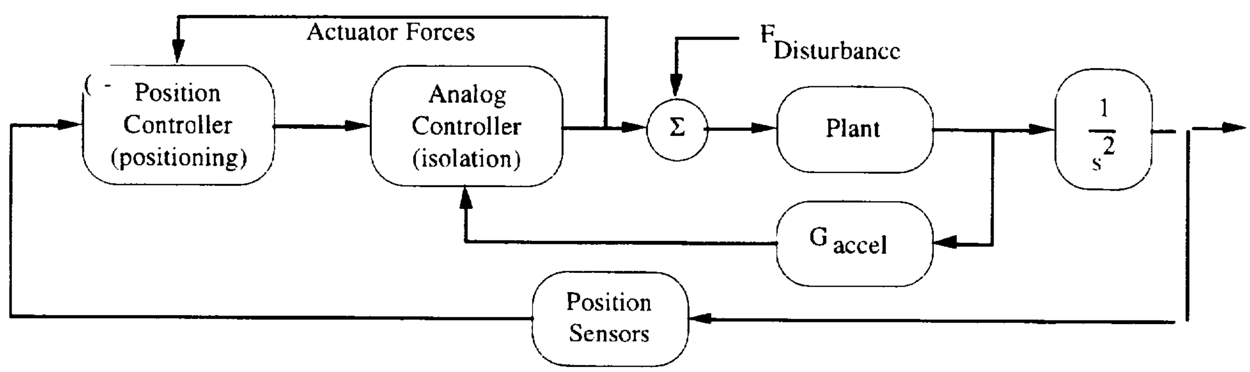


Figure 3.- STABLE Control System Architecture

The electronics that controlled the STABLE system was split into four cards: digital control, analog control, actuator driver and power supply. The core processor for the STABLE digital position controller was an INTEL 80C196KC16-bit microcontroller operating at 8MHz. The microcontroller's analog-to-digital converter (ADC) was operated in the 10-bit mode. The inputs to the ADC were the position sensor outputs, current feedback signals, and the accelerometer temperature measurements. The outputs from the position controller were sent to 12-bit digital-to-analog converters (DAC) on the analog control card and summed with the accelerometer feedback. This signal was passed through a 12-bit multiplying DAC. The multiplying DAC allowed the position controller to take command of the overall system in high-gain mode and provided the flexibility to change the sign of the feedback signal. The combined signal was then sent to the analog control electronics which incorporated a low-pass filter stage with a nominal bandwidth of 50Hz. This results in an acceleration or current command signal. The actuator driver card amplified the current signals using a linear power op-amp stage which provided the actuator drive current.

A data acquisition system (DAS) recorded the STABLE system parameters generated by the control electronics and several temperatures from the CHUCK experiment. The CHUCK video signals were recorded on an 8mm camcorder external to the STABLE system. The parameters measured were: six platform accelerations, three base accelerations, six current commands, twelve position signals, four CHUCK experiment temperatures, and six mode bits. These parameters were recorded by two PC-cards inserted into a Payload General Support Computer (PGSC-486), which is a flight upgraded version of an IBM 750C Thinkpad computer. The PC-cards were 16 channel

12-bit ADC cards with 8 bits of digital I/0 from ComputerBoards Inc.. In addition to the previous parameters, nine accelerometer temperatures and a program overflow byte were sent via RS232 to the PGSC-486. The data acquisition system sampled and recorded the accelerometer and current parameters at 250Hz and all others at 10Hz. The DAS system stored the data on the PGSC-486 500-Mbyte removable harddisks. Eight of these harddisks were used in flight and about 72 hours or 3.5-Gbytes of data was recorded during the mission.

The STABLE system was designed for minimal crew interface and attention. The main tasks of the crew were to setup the PGSC-486, enter the current Mission Elapsed Time (MET) and activate STABLE. STABLE utilized six manual switches located on the front panel with corresponding i lluminating indicators. These switches were for main power, position mode, acceleration mode, fan power, CHUCK power, and CHUCK heater power. When position mode switch was activated, STABLE would operate only in high-gain mode. When both the position switch and acceleration mode switches were activated together, the STABLE system would operate in the nominal low-gain mode. Other activities performed by the astronaut crew were: change out full PGSC-486 harddisks with a replacement harddisk and operate the CHUCK experiment. CHUCK was operated only four times during the mission, for a total of 8 hours, and its diffraction images recorded on videotape for post-mission evaluation.

The STABLE system achieved its objectives by isolating the CHUCK microgravity-sensitive payload from the "noisy" USML-2 environment and served as a microgravity isolation technology demonstration. Several preliminary analyses of the STABLE data have been performed by MSFC, LeRC, and MDA, and more detailed analyses are under way. Generally, the time history data shows a 25 times reduction in the level of disturbances reaching the isolated platform from the external environment. This isolation factor is consistent throughout the frequencies measured: 0.02Hz - 125 Hz. Another performance measurement is the power required by the STABLE system to perform the isolation. The nominal power required by the actuators during a thruster event in which STABLE maintained isolation was about 3 Watts. This power is in addition to the overhead electronic component usage of about 25 Watts.

CURRENT ISOLATION TECHNOLOGY AND FUTURE ADVANCES

This section describes the current technology behind the other systems mentioned in the introduction and gives the reader an idea of the future direction and recent advances in isolation technology. Since some of these systems, MIM and ARIS, have not disclosed data received from their flight experiments, no performance estimate of these systems is included. For an in-depth analysis of any mentioned technologies and isolation systems, the reader is forwarded to the references on these systems from the original developer.

Generally, it can be stated that the current direction of isolation technology development is to miniaturize the hardware components while maintaining the accuracy, stability and performance of the overall system, The isolation technology hardware descriptions will be separated into the following sections: actuators, sensors, control electronics, and umbilicals, Vibration isolation technology is currently being developed at a rapid pace and engineers are continually developing new methods to isolate systems from noisy environments. Since new technology advancements from other developers are not currently known, only those advances that pertain to future STABLE type systems are included.

ACTUATORS

While the STABLE system used wide gap MDA electro-magnetic actuators, other isolation systems use many different types of actuators, including mechanical, electro-magnetic, and electro-mechanical. The MIM and the LeRC systems use a Lorentz force electro-magnetic type actuator. In their actuators, the area of the magnet is large compared to the coil area, which is different from the MDA actuator, where the magnet area is small compared to the coil area. The ARIS system uses a small-angle motor which drives thin pushrods attached to isolated payload. In this configuration, the ARIS system has a direct mechanical linkage to the payload. Honeywell has used various types of actuators in their systems. Their FEAMIS system utilized electro-magnetic actuators in which an electro-magnet attracts a metallic plate. To isolate the Hubble Space Telescope Reaction Wheel Assemblies, Honeywell used a hydraulic isolator similar to a shock absorber. Some isolation systems, such as 6-DOF STEWART platforms, use traditional hydraulic actuators, linear motors or electric motors attached to screw type mechanisms.

Except for the wide gap, which is necessary to accommodate the required rattlespace, the design of an electro-magnetic actuator is similar to an electric motor, and it is theorized that electric motor technology could be adapted to microgravity actuator design. Exotic **magnetic** materials, such as samarium-cobalt and neodymium, could be used to develop stronger permanent magnetic fields. Advanced methods to contain the stray magnetic field can be implemented to develop more compact and powerful actuators. The goal in future actuator designs is to make the actuator size small, have minimal amount of rattlespace for the projected environment and utilize very little power to provide the required forces. An actuator design that meets these objectives requires a study of the following: magnet ic materials and flux design, wire size, number of windings, coil orientation and collocation of sensors.

SENSORS

Isolation systems use many different types of sensors for determining the position of the floated platform with respect to the base. Some of sensor types used are: photovoltaic, photodiode, hall-effect, capacitive, linear transducers, optical encoders, resolvers and CCD cameras. Both the STABLE and MIM systems used the dual-axis photodector and laser combination while the ARIS actuators have an internal single-axis optical position sensor with a light-emitting-diode (LED) for the source. The LeRC system used a permanent magnet with hall-effect sensor combination and the Honeywell FEAMIS system used a hall-effect sensor for controlling the actuator's magnetic gap. STEWART platform motors and hydraulic actuators could utilize any conventional type of position sensing device, resolver, encoder, or other linear transducers.

Collocation of sensors and actuators is the current trend in the design of STABLE type isolation systems. The position sensors need to be non-contacting and be able to measure the full system rattlespace range. Also, the position sensors need to be insensitive to magnetic fields if located near an electro-magnetic actuator. Unfortunately, photodectors are sensitive to magnetic fields and need to be magnetically shielded or located away from electro-magnetic actuators. A new position sensor concept in development is utilizing the inductance change of the actuator armature coil relative to the magnet position. This unique idea inherently collocates the actuator and position sensor and would not require many external electronic components.

Measuring accelerations in the micro-g range, 1×10^{-6}g or µg, is not simple. The Sunstrand QA-2000 accelerometer was used by STABLE, LeRC, MIM and also by the Space Acceleration Measurement System (SAMS). The QA-2000 series uses an electro-magnetic coil for force rebalancing and outputs a current signal proportional to the measured acceleration. Unfortunately, some problems exist with this accelerometer: liftoff forces can change the bias, the coil is sensitive to magnetic fields, and the sensitivity to temperature changes is significant at low acceleration levels. Sunstrand has addressed the temperature sensitivity problems in their new QA-3000 series which was used by the ARIS system. Also, depending upon the expected microgravity environment, the sensor might have a dynamic range of 0.1 µg to 10,000 pg. This dynamic range is significant for the sensor, amplification stages, and control system. Finally, the noise floor of an acceleration measurement system must be considered. Accelerometer noise can be attributed to aliasing and quantization effects from digital sampling and the manufacturer's accelerometer noise specifications.

Again, collocation of the accelerometer with the corresponding actuator's axis is important. This reduces the effects of rotational accelerations from corrupting the actual acceleration measurement. Current developments in accelerometer technology include the use of micro-machining to develop sensors that can be placed on an integrated circuit along with the compensation electronics. This advancement will yield sensors that are lighter, more sensitive, and use less power than current technology. Several companies, Endevco, Clifton Precision and EG&G [C Sensors, are working on these types of accelerometer designs for automobile airbags, and it is theorized that these sensors could accommodate microgravity acceleration ranges by changing the internal coefficients.

CONTROL ELECTRONICS

The control systems and electronics for isolating payloads are as varied as the scientific payloads themselves. STABLE is unique in that it utilizes a combination of both analog and digital control. The MIM and ARIS systems use only digital control for both the acceleration and position control laws and both use high performance digital signal processors (DSP's) for performing the control law calculations. The ARIS system employed the 3 Dimensional Microgravity Accelerometer (3-DMA) system to perform overhead functions and communicated with 3-DMA using both RS232 and MIL-STD-1553B data links. The 3-DMA overhead functions are: recording acceleration measurements, interfacing to the PGSC-486, change control gains, reprogramming, postprocessing, command, control, and data uplink and downlink. To perform its overhead functions, the MIM system incorporated a 486 type single-board computer into its electronics package. These high performance processors, DSP's and 486 computers, compute the control law equations, transformation matrices, sample sensors, and perform overhead functions at a very fast sampling rate, usually ~1000Hz. This fast sampling rate is necessary to achieve a reasonable control bandwidth of- 100Hz. Unfortunately, these high performance processors are currently very susceptible to single-event upsets (SEU's), which can cause systems to go to an unknown state. If the SEU's are not accounted for by software or hardware means, they can cause an isolation system to "crash" into the containment bumpers, thus violating the isolation requirements. To overcome this problem, the ARIS and MIM systems have software and hardware watchdog timers that reset the system in case a SEU was detected.

One of the advances in digital control will actually come from the collocation of position sensors and actuators. Collocation will eliminate the need to transform the position information into actuator gap coordinates, allowing multiple 2-DOF systems to operate together to achieve isolation. Reducing the calculation time, by collocation, will not reduce the trend to use DSP's and 32 or 64-bit processors. Since these high performance processors are designed for these applications, the research direction will be to find means to either shield or radiation-harden the

core processor from the space environment. Once these processors become space qualified, the use of both analog and digital control together will decrease. If the position control bandwidth in isolation systems continues to remain relatively low, ≤ 1 Hz, the use of slower, more accurate, >16-bit, ADCs will become prevalent, allowing precision positioning of the isolated payload. Another advance in data transmission for isolation systems will be the use of either infrared (IR), radio-frequency (RF) or optic fiber technologies. New advances in cellular phone technology and the new Infrared Data Association (IRDA) standard 1.1 will allow for transmission of data between the payload and base, at rates required for control of payload physical processes. The transmission interface will require a minimal set of overhead electronics, but this transmission method would reduce the number of umbilicals.

UMBILICALS

Flexible umbilicals are critical in the design of isolation systems. The stiffness of the umbilicals is the primary factor in determining the required actuator forces, the transmission of disturbances to the payload and the control system bandwidth. The umbilicals must also have the capacity to transfer resources to and from the payload. These resources can include any combination of: power, data, video, vacuum resource and exhaust, fire detection and maintenance, coolant supply and return and gaseous supply. The STABLE experiment used two flex ible umbilicals, one for power and the other for data and video, while the MIM system provided only power and data services to the isolated experiment. The ARIS design included a full complement of umbilicals and was the first isolation system to integrate many different types of umbilicals and materials in the design process.

Recent advances in umbilical technology include: umbilical material studies, coiling, splints, and umbilical followers. A study of new materials and umbilical designs is being performed for the Space Station Furnace Facility project by the MSFC Propulsion Laboratory. A Dupont/Dow F200 barrier hose material is one of the materials being tested for vacuum, fluid and inert gas umbilicals. The materials and manufacturing processes for all flight umbilicals are defined by the Space Station Qualified (SSQ) specification. Even though these materials have a low stiffness property, it is also helpful to coil the umbilicals to further reduce the stiffness. Figure 4 shows a set of coiled umbilicals for the ARIS system, If two umbilicals touch, the disturbance transmission path is "shorted," which increases the effective stiffness. To prevent this from occurring, small splints can be inserted to keep adjacent coils from touching. The splints can be semi-flexible to help isolate adjacent umbilicals, but rigid enough to keep the coils at a predetermined distance. Although these methods reduce umbilical stiffness, large bias forces can develop from the movement of the isolated payload. To overcome this problem, an umbilical follower can be implemented to reduce these bias forces. An umbilical follower is a 3-DOF mechanism that moves the base of the umbilical set to minimize the average actuators' currents, which corresponds to the umbilical bias forces.

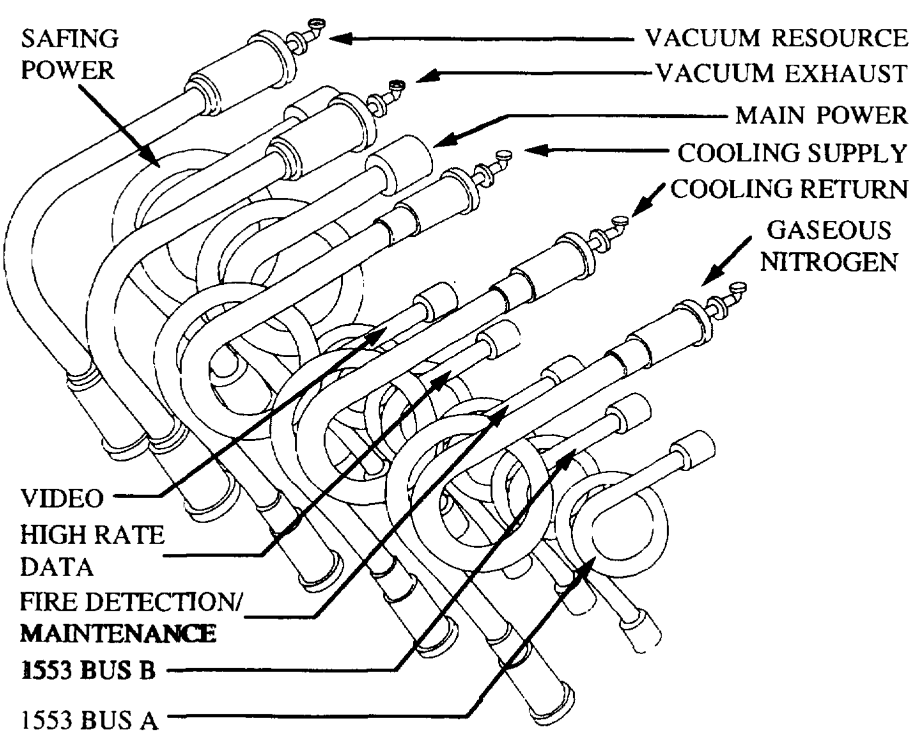


Figure 4.- ARIS System Umbilicals

CONCLUSIONS AND FUTURE DIRECTIONS

The microgravity isolation field has a long history beginning in the early 1970's with the Honeywell corporation. Twenty-five years later, isolation technology has improved significantly and the need for microgravity isolation is now more important considering the many different types of disturbance sources, known and unknown, present on the ISS. The STABLE system was the first flight of an experiment to demonstrate microgravity isolation and was soon followed by the MIM and ARIS systems. Although MIM and ARIS have not disclosed their isolation performances, the STABLE system proved that experiments can be isolated from space environment disturbances, including crew activity, while using minimal resources. All the isolation systems mentioned have assisted in defining the core technology for microgravity isolators. The technological basics of implementing microgravity isolation are very straightforward, and anyone skilled in motion control technology can make significant contributions. Any contributions to the core technologies, actuators, sensors, umbilicals, control systems and electronics, may apply to many different industries with similar requirements. Some of these industries include motion control, robotics, vibration suppression, manufacturing and vehic le ride control. Core technological research, like MSFC's umbilical material and design study, is essential for developing the next generation of hardware components for microgravity isolators.

Current technology development and research are directed at every aspect of isolation systems. New types of actuators, sensors, umbilicals, control architectures and electronics are being developed as well as a new isolator design for ISS racks, experimental payloads and the ISS microgravity glovebox. Future actuator designs will be smal ler with improved magnetic flux containment and have both position and accelerometer sensors collocated with the actuator. Micro-machined accelerometers specifically designed for microgravity applications will be necessary for smaller isolator designs. Novel position sensor techniques are being developed which integrate the actuator coils as a position sensor. New umbilical materials and umbilical followers are being designed to reduce bias forces. To control isolator systems, future designers will use space qualified microprocessors and DSP's and more accurate, ≥16-bit, ADC's and DAC's. Current control architectures use either digital or a combination of digital and analog control, but future isolators will include some type of adaptive control to account for different payload inertial properties. Finally, new isolator designs might use either IR, RF or even optical fibers for data transmission between the base and the isolated platform. In general, new and improved isolation systems need to be designed which have simple interfaces and use minimal overhead volume and power. In summary, these current areas of study and other technologies yet to be invented will advance isolation systems into the twenty-first century.

REFERENCES

1. Wolke, P. J.: "Advanced Technology for Active and Passive Vibration Isolation of Spaceborne Payloads," Pub, No. S69-5504- 1.0-1, NASA Vibration Isolation Technology for Microgravity Science, Cleveland, OH, September 1988,

2,. Havenhill, D., Wolke, P. J.: "Magnetic Suspension for Space Applications." Pub. No. S69-5 104-2.0-0, Honeywell Inc. Report, Glendale, AZ, October 1990.

3. Grodsinsky, C. M.: "Microgrvity Vibration Isolation Technology: Development to Demonstration," NASA TM-106320, September 1993.

4. Lumboski, J.F, et al.: "Final Report-Vibration Isolation Technology (VIT) ATD Project." NASA T.M-106496, March 1994.

5. Boeing Defense & Space Group, "Prime Item Development Specification for the Active Rack Isolation System." NASA Specification Number S684- 10158, June 1995.

6. Tryggvason, B.V., Salcudean, S. E., Stewart, B. Y., Parker, N.: "Microgravity Vibration Isolation Mount (MINI)." Presentation at 12th MGMG Meeting CSA, Cleveland, OH, May 1994.

7. Edberg, D., Boucher, R., Schenck, D., Nurre, G., Whorton, M., Kim, Y., Alhorn, D.: "Results of the STABLE Microgravity Vibration Isolation Flight Experiment ." A AS Paper 96-071, American Astronautical Society Conference, Reno, Nevada, February 1996.

Page intentionally left blank

URC97008

An Adaptive Learning Strategy for Autonomous Machines

G.S. Alijani, A. Omar, J.S. Welsh
Computer Science Department
Southern University at New Orleans
New Orleans, LA 70126
dalijani@ix.netcom.com

Abstract

This paper presents the design and evaluation of an adaptive learning strategy for robotics systems capable of performing time-critical missions. The strategy is a combination of a *case-based* learning method and learning by *instruction.* The learning method was implemented in three phases: *training, monitoring,* and *autonomy.* In addition, a hard real-time scheduling scheme was provided to schedule critical and non-critical tasks. Measurements were obtained in terms of processing time, number of trained sequences, number of matched sequences, and the number of completed sequences with deadlines,

1. Introduction

The objective of machine-learning is to develop a machine capable of accepting a complex real-world problem and reaching a correct solution by executing the correct sequence of primitive tasks. Two challenging parts of developing a learning machine are the choice of a knowledge representation scheme and the transfer of logical functions to the machine. For the last two decades, researchers have developed a variety of learning strategies, including *incremental* learning [1], case-based and explanation-based learning [2], and learning by *observing* [3]. In some of these methods, a trainer provides the learner with basic information or feedback, while others depend upon the learner to seek or "scl f-organize" the information.

Learning by *instruction* that is used in programming manipulator sequences. In this method, the learner accepts general concepts or facts from a trainer and then repeated] y applies them to specific instances [4]. While this approach is easy to implement, it is inflexible and the correctness of the sequence relies solely on the instructor. Learning by *example is* another common method used for robotics systems [5]. The cased-based learning is a special case of learning by example in which the nearest match between the current state and a stored state is used to generalize a plan and predict the possible results [6].

Neural networks offer a new approach to train a programmable system to perform selected tasks or solve complex problems. An intriguing feature of neural networks is their ability to continue to function even when some of its processing elements fail [7]. This characteristic makes the neural networks more suitable for critical missions and real-time control systems [8]. However, neural networks require more computation power than any other method. More recent learning methods are being developed based upon applications of Fuzzy logic [9, 10].

2. Learning Strategy

It has been shown that no single technique is adequate to build a complete learning system [2]. Consequently, attempts have been made to combine two or more different approaches. An example of a combinational strategy is the adaptation-based explanation, which brings together some of the better insights of both case-based and explanation-based learning methods.

The new learning strategy called *adaptive learning,* that is proposed and tested here, is a combination of learning by *instruction* and *case-based* learning methods. Learning by *instruction* is an excellent paradigm for setting up the initial knowledge, especially when the system is unable to handle an unpredictable situation. Then, when the robotics system is operating, there arc a significant number of periodic tasks and it is highly probable that a new task will be similar to a past task, making the case-based method an efficient strategy. The combination of these two methods should meet most requirements for developing an adaptive learning strategy.

Within the adaptive learning strategy, a cycle of *training, monitoring,* and *autonomy* phases are defined. In the training phase, learning by *instruction was* applied to train and monitor the system to perform a set of primitive tasks, PT_i's. The system was then directed to perform an *ordered* sequence of primitive tasks, defined as a subtask, ST_j. Once al 1 the subtasks were completed, the trainer provided the robot with the logical functions associated with each PT_i. The robot stored the tasks, along with their associated functions and data. It was expected that, using this stored knowledge, the robot would be able to perform the same or a similar sequence of tasks independent y,

The objective of the monitoring phase is to test and improve the efficiency of the robot's performance. During this phase, the trainer asks the robot to perform a previously defined task. Since each complete task consists of a sequence of ordered subtasks, the robot has to search through its memory to generate a global execution plan. If an exact match does not exist, then the robot has to make a decision based on similar past cases to develop an execution plan. Finally, if the robot is unable to find either an exact match or a similar case, it asks the trainer for a global execution plan. Thus, the ability of the robot to find a solution could be measured by (he number of its requests to the trainer.

In the autonomy phase the robot is allowed to make decisions by searching past experiences and executing a gl obal plan using the logical functions associated with each subtask. If the robot tries al 1 the alternative sequences and it cannot find a task that exactly or partially matches the requested task, it will ask the trainer for a global execution plan. Thus, in this phase, only in an extreme case, in which the robot has absolutely no background on an assigned complete task, it will ask the trainer for the execution plan.

Real time training of the robot involves four basic steps; (1) the robot collects information on the operating environment; (2) the trainer develops an execution plan (by manipulating collected data and using logical functions) and sends the plan to the robot; (3) the robot executes the plan and sends the result back to trainer, and (4) after the trainer confirms the correctness of the execution, he sends the data and the logical function back to be stored in robot memory. After the robot learns about all the PT_i's, two or more PT_i's are grouped into a *unique ordered sequence* to establish a subtask, ST_j. The same procedure *is* applied to group a sequence of ST_j's into a complete task, CT_k. Complete tasks are classified into two groups: *periodic* and *sporadic.* The periodic tasks are predictable, with a fixed ordered sequence of subtasks, while the sporadic tasks are unpredictable and the number of subtasks can be changed during execution. To handle the two different classes, allowances are made for both *static* and *dynamic* grouping in which the number of subtasks can either be fixed or varied.

In addition to the learning strategy, further enhancements are needed to manage the information and schedule activities. Our system is supported by a *memory manager* and a real-time *task scheduler,* as described in the following sections.

3. Memory Manager

The purpose of the memory manager is to organize classes of tasks (primitive, subtasks, and complete tasks), logical functions, and associated data so that the scheduler can readily access and

schedule tasks for execution. Initially, the robot starts with an empty memory. As the training procedure progresses, tasks, logical functions, and the required data are sent to the robot. Eventually, the robot will have the same logical memory structure as the control station (trainer). Figure 1 depicts the primitive task, subtask, and complete task lists as they are linked together in a hierarchical fashion.

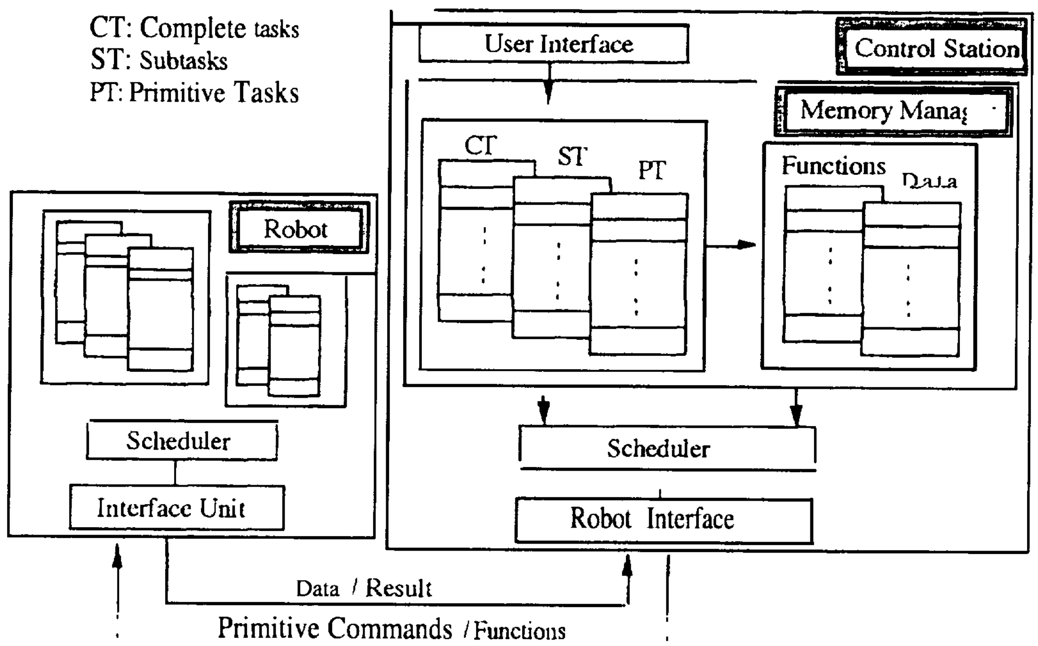


Figure 1 : Logical Structure

Each CT_k is linked to a list of related ST_j's, where each ST_j in turn links to a list of PT_i's. Each primitive task is linked to a specific function and an associated data set. Further, the list of complete tasks is divided into two lists: periodic tasks, P.list, and sporadic tasks, S.list. To facilitate searching and manipulating procedures, the P.list is arranged using *a* Most Frequently Used (MFU) technique, while the S.list utilizes a Most Recently Used (MRU) mechanism.

4. Task Scheduler

A task scheduler is another essential element in a self-content system that operates in a real time environment. Generally, the tasks scheduler manages the resources and schedules critical and non-critical tasks according to their deadlines. In our case, as Figure 1 illustrates, both the robot and the control station are provided with a task scheduler. According y, the robot and the control station can manage their own resources and schedule their own activities. In conjunction with the memory manager, the tasks scheduler determines the correct sequences of PT_i's, collects all the functions and the associated data, produces a sequence of ST_j's or a CT_k, and schedules them for execution.

In the last two decades, extensive research has been conducted in the areas of task scheduling [11, 12], response lime [13] and resource management [14]. The task scheduler used in this paper is a three-phase task scheduler that was designed based on *commitment* and *look-ahead* strategies. A detailed description of these strategies and a performance evaluation of the task scheduler can be found in [11, 15].

5. Performance Evaluation

The following criteria comprise the guidelines for evaluating the model's performance:

- The number of trained sequences versus the number of requested sequences (randomly generated by the trainer and requested for execution),
- The number of trained sequences versus the processing time,
- The number of trained sequences versus the number of completed sequences" with deadlines, and
- The processing time versus the depth limit and the threshold value.

Each category was tested when the robot was functioning in the monitoring and the autonomy phases. The processing time is defined as the time that a robot needs to match a similar past case with the requested task.

To determine the ability of the robot to match requested sequences, a series of experiments was conducted. In each experiment, a robot was trained for a sequence of well-defined tasks, then it was asked to match a sequence of *randomly* generated tasks and to develop execution plans. As Figure 2 indicates, the robot is capable of matching 87 percent of the requested sequences. Since the requested sequences are generated randomly, the order of tasks within some of the generated sequences are very likely to be different from those within any of the trained sequences, perhaps explaining the missing 13 percent. (There are also some unpredictable tasks which the robot has never trained for.) A final issue is the trade-off between processing time and the number of the matched sequences. These experiments provide insights into the selection of a feasible solution in which the robot can provide both the maximum matched sequences and also satisfy the time constraints.

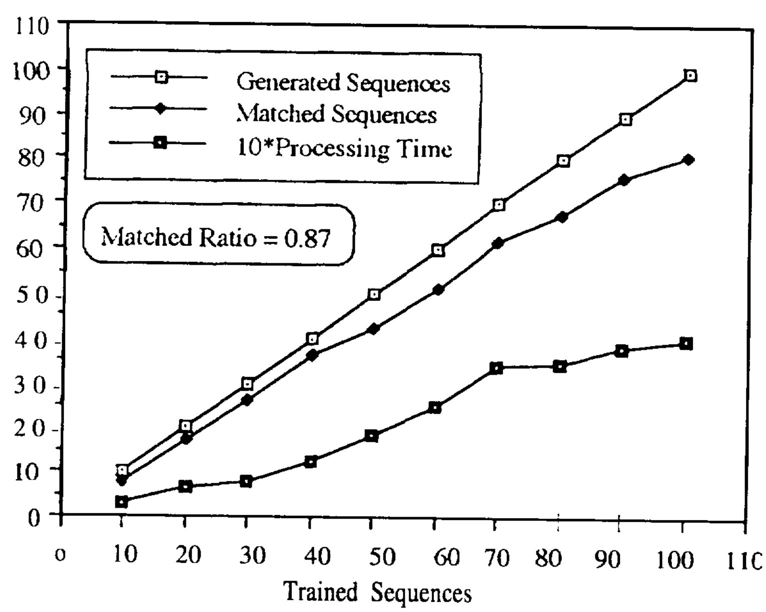


Figure 2: Trained Sequences vs. Matched Sequences and Processing Time

Since the system is operating in a real-time environment, considerable time is needed to match sequences generate execution plans, and schedule the critical and non-critical tasks. Thus, it is essential to determine the cause of deadline failures, if any. A deadline failure may be caused by the inability of the robot to execute requested tasks, a deficiency in the memory manager or the scheduler, or a combination of factors. To determine the capability of the system in completing sequences and meeting their deadlines, another series of experiments was conducted. As Figure 3 indicates, up to 70 sequences, the system can meet 90 percent of the requested sequence deadlines. However, the

performance of the system degraded once the number of requested sequences approached 100, possibly due to excessive processing time.

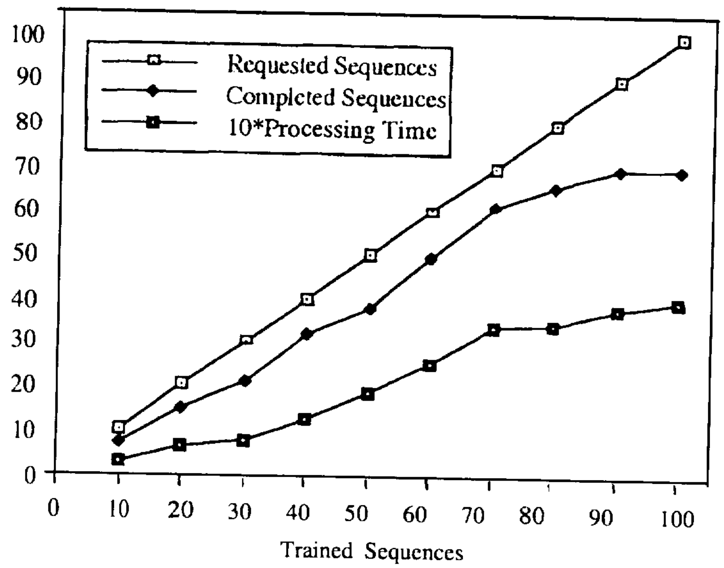


Figure 3: Requested Sequences vs. Completed Sequences (meet their deadlines)

To investigate the effect of searching and execution times (processing time) on meeting the deadlines of requested sequences, depth *limits* and *threshold* values are defined. The depth limit determines how far the list of subtasks should be searched to find a exact match to a subtask. A threshold value is the percentage of subtasks within a complete task that should be matched to a similar past case. To satisfy a threshold value and a depth limit, the memory manager must search through its hierarchical memory structure (Figure 1) and select an ordered sequence of sub tasks using the MFU or MRU schemes. The effect of these two factors on the processing time is shown in Figure 4. These experiments provide insights into both the selection of partial or exact matching of subtasks and meeting the deadlines of complete tasks.

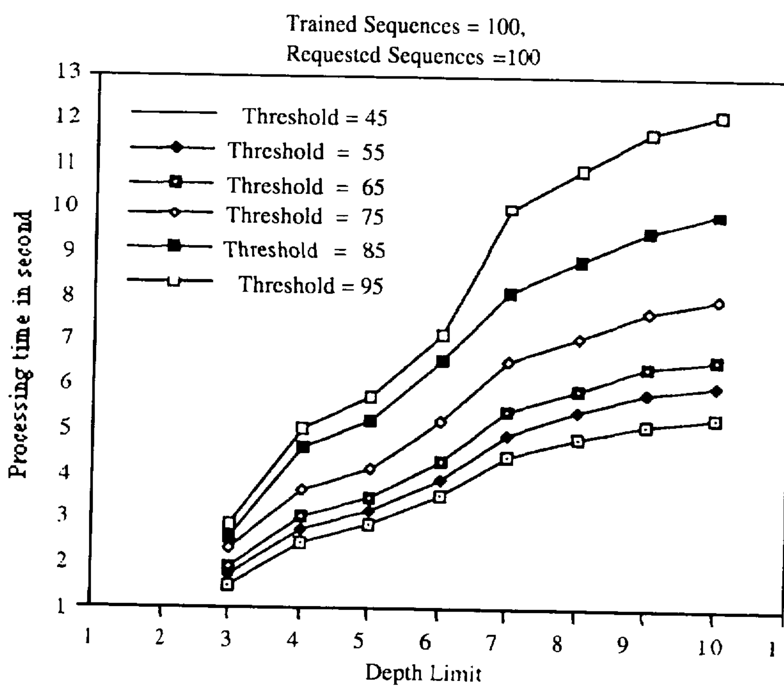


Figure 4: Depth Limit and 'threshold Values vs. Processing Time

6. Conclusion

Learning by *instruction* and a *case-based* method were combined to define an adaptive learning strategy in which a robot was trained and monitored on a real-time basis. The system was supported by a hard real-time task scheduler and was evaluated in terms of processing time, the number of both trained and matched sequences, and the number of tasks that met their deadlines. This research contributes to the areas of real-time systems and autonomous machines. It provides insights to the design and evaluation of autonomous systems by utilizing a potentially powerful knowledge representation method, an adaptive learning strategy, and a red-time task scheduling scheme.

7. Acknowledgments

This research was supported in part by the Office of Grants and Sponsored Research at the Southern University at New Orleans. We also like to thank Gina Jackson for her constructive comments.

References

[1] J.J. Grefenstette, "Incremental Learning of Control Strategies with Genetic Algorithms," *Proc. of Sixth Int. Workshop on Machine Learning,* 1989, pp. 340-344.

[2] M. Redmond, "Combining Case-Reasoning, Explanation-Based Learning, and Learning from Instruction, "*Proc. of Sixth Int. Workshop on Machine Learning, 1989,* pp. 20-22.

[3] D. Wilkins, B. Buchanan, and W. Clancey, "Inferring an Expert's Reasoning by Watching, "*Proc. of the* 1984 *Conf. on Intelligent Systems and Machine*, 1985. pp. *84-92*.

[4] K. R. Levi, V.L. Shalin, and D.L. Perschebacher, "Learning plans for an Intelligent Assistant by Observing User Behavior," *Int. Journal of Man-Machine Studies*, 1990, pp. 489-503.

[5] C.M. Kadi, "Diffy-S: Learning Robot Operator Schemata from Exam pies," *Proc. of the Fifth Int. Conf. on Machine Learning,* 1988, pp. 430-436.

[6] C. Sammut and D. Hume, "Observation and Generalization in a Simulated Robot World," *Proc. of the Fourth Int. Workshop on Machine Learning*, 1987, pp. 267-273.

[7] M. Gaudill, "Neural Networks Primer, Part I," *AI Expert,* December 1987, pp. 46-52.

[8] H. Klopf, "Neural Network May Make Smart Robotics," *Machine Design,* June 8, 1989.

[9] A. Matinez, E. Tunstel, and M. Jamshidi, "Fuzzy Logic based Collision Avoidance for a Mobile Robor," *Proc. of Int. Conf. on Industrial Fuzzy Control and Intelligent Systems,* pp. 66-69, 1993.

[10] C-U Wang, T-P Hong, and S-S Teng, "Inductive Learning From Fuzzy Examples," *Proc. of the Fifth IEEE Int. Conf.* on *Fuzzy Systems,* pp. 13-18, 1996.

[11] G.S. Alijani and S.C. Su, "A Real-Time Task Scheduling Scheme using Loosely Coupled Systems," *Proc. of the third Conf. of the North America Transputer User Group,* pp. 128-137, April 1990.

[12] J. Stankovic and K. Ramamamritham, "A Reflexive Architecture for Real-Time Operating Systems, " Chapter in *Advances in Hard Real-Time Systems,* Prentice Hal 1, pp. 23-38, 1995.

[13] W.C. Wesley, C.M. Sit, and K.K. Leung, "Task Response Time For Real-Time Distributed Systems With Resource Contentions," *IEEE Transactions on Software Engineering, Vol. 17, No. 10,* October 1991.

[14] K.G. Shin, C.M. Krishna, and Y .H. Lee, "Optimal Resource Control in Periodic Real-Time Environments," *Proc. Real-Time Sys. Symposium,* December 1988.

[15] G.S. Alijani and et al., "Schedulability: Real-Time Tasks Scheduling in Mutliprocessing Systems," *Proc. of the 23rd Int. Conf. On Parallel Processing,* Vol. 11 I, pp. 79-82, 1994.

Quantifying Multirate Control System Component Sample Rate Change Performance Effects

Robert J. Alvarez
Department of Mechanical Engineering
University of New Mexico

Michael A. Kvasnak
SVS, Inc.
Albuquerque, NM

Mohsen Shahinpoor
Department of Mechanical Engineering
University of New Mexico

Abstract

Current multirate control design methods do not provide the control system designer with the rationale to choose individual component sample rates; seemingly insignificant changes in those sample rates can, however, produce significant changes in system performance. This paper introduces the notion of effective latency -- the time it takes data to move from the sensor, through the A/D converter, compensator, and D/A converter, to the plant -- and shows that it is possible to use effective latency to predict the performance of a multirate system. The mean effective latency is shown to be a predictor of step response stability; jumps in mean effective latency can also be used to signal major changes in multirate system behavior.

Introduction

This paper describes preliminary results in the search for an algorithm to predict multirate control system performance as a function of individual component sample rates. The term multirate is used in a strict sense to represent a system where the individual sampled-data devices -- analog-to-digital (A/D) converter, compensator, and digital-to-analog (D/A) converter -- are run by independent processors operating at different, but constant, sample rates. The term synchronous is used to describe a subset of multirate systems where all sampled-data devices run simultaneously at equivalent sample rates. The block diagram representation of an example multirate control system is presented in Figure 1. The sampled-data devices in the example are shown in italicized font; though most multirate control design approaches constrain the individual sample rates to be integer multiples of each other [1], there is no such implicit requirement here.

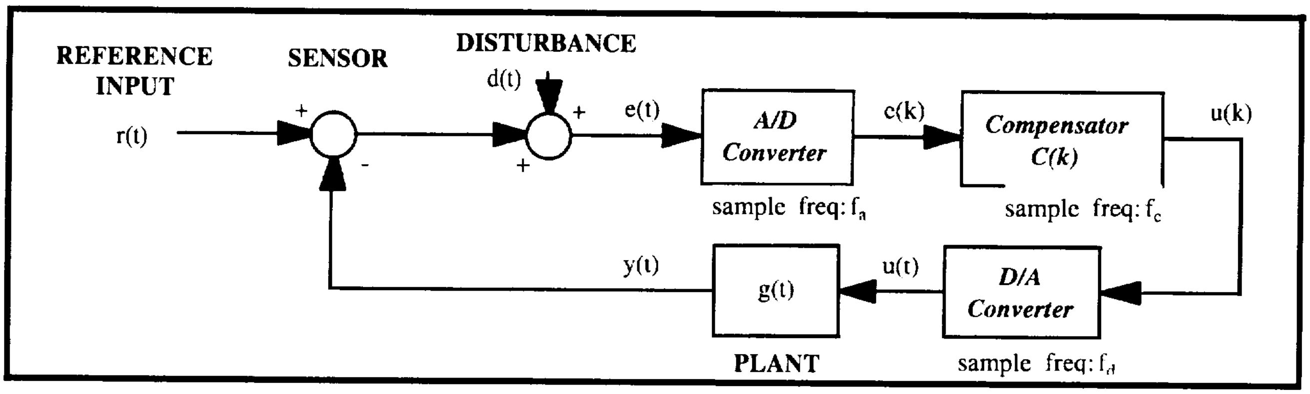


Figure 1. Multirate Control System Block Diagram

Current multirate control design methods do not provide the designer with the rationale to choose individual sample rates [2]. There are methods, based upon extended state transition matrices, to predict multirate system stability [3], but predicting multirate system performance metrics (such as rise time and mean-squared error) in light of changing individual component sample rates has proved very difficult [3]. Insignificant changes in individual component sample rates can, without warning, produce significant changes in system performance [4], Figure 2 shows an example of the counter-instinctive changes in plant output step response due to changes in D/A sample rate; as the D/A sample rate increases from 19.8 kHz to 28.3 kHz, the plant output step response peak magnitude increases from 1.5 units to 2.1 units, then decreases to 1.25 units, then increases again to 1.85 units. (The step responses were simulated using an event-based simulation which updates the system state after each sample time. System parameters used as the basis for the step responses are included to the right of the step response plots. For this example, the sampled-data devices are assumed to have no effect on the magnitude of the input signal, but are instead modeled simply as sample-and-hold elements.)

The objective of this paper is to show that, though the changes in system performance as a function of individual component sample rates may seem to be unpredictable, there are indeed clues which make this prediction possible. By introducing the notion of effective latency -- the time it takes data to move from the sensor, through the A/D converter, compensator, and D/A converter, to the plant -- it is actually possible to predict the performance of a multirate system based solely upon its individual component sample rates. This type of knowledge will allow the control system designer to "tune" the performance of a multirate control system by simply adjusting sampled-data device sample rates. The next sections of this paper will present an examination of the performance effects induced by such adjustments.

Sample Rate Changes in a Synchronous System

Since the multirate control system performance changes induced by changes in individual component sample rates arc difficult to predict, it is best to first demonstrate the applicability of effective latency in the prediction of synchronous system performance. To create a synchronous system, it is assumed that the designer can simply rein in the faster sampled-data devices to force all components to begin new tasks at the same time.

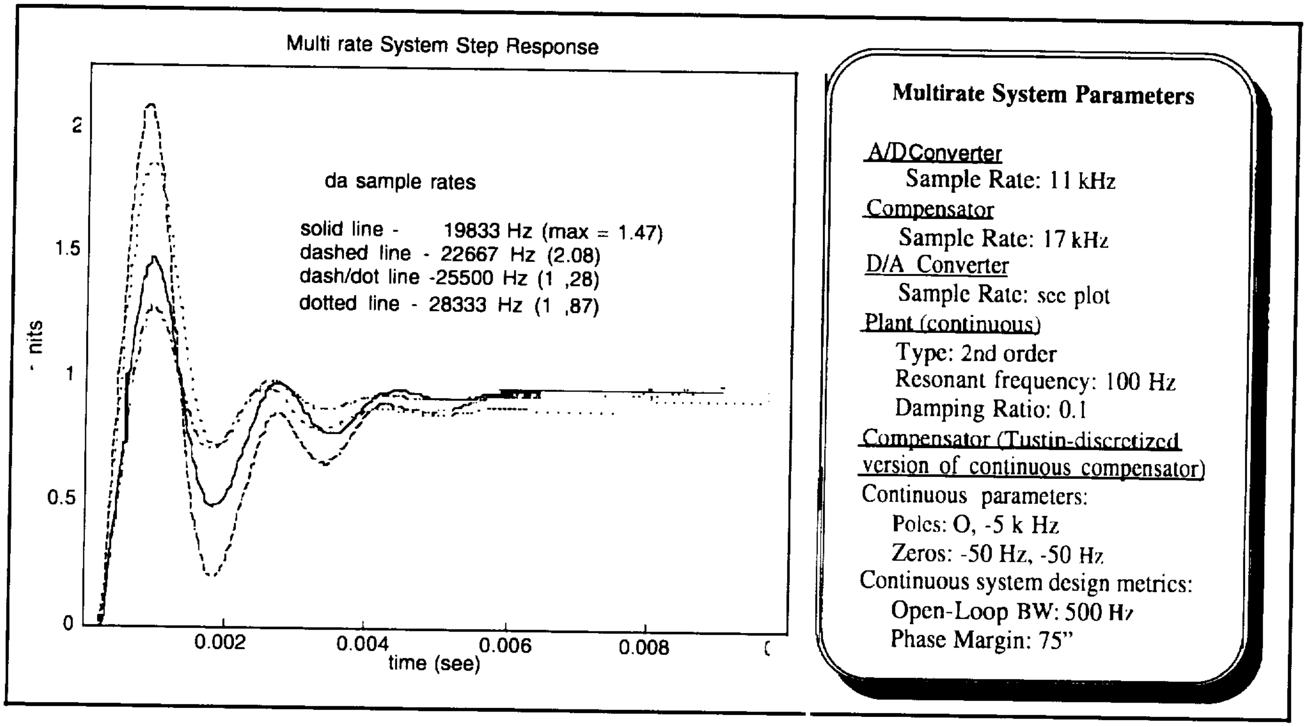


Figure 2. Multirate System Step Responses and Parameters

The left side of Figure 3 shows a cartoon example of the timing between individual component events in a synchronous system; the times t_{ai}, t_{ci}, and t_{di} represent starting times for the A/D conversion, compensation calculation, and D/A application tasks. The bottom of the figure presents the effective latency: a single sensed error measurement enters the sampled data portion of the closed-loop system at t_{a1}; the digitized representation of that measurement is input to the compensator at t_{c2}; the compensated command is transferred to the D/A converter at t_{d3}; and, the analog equivalent to that command is applied to the plant upon completion of the D/A task. The effective latency of that initial sensed error measurement is the sum of all those times. This latency remains constant in a synchronous system.

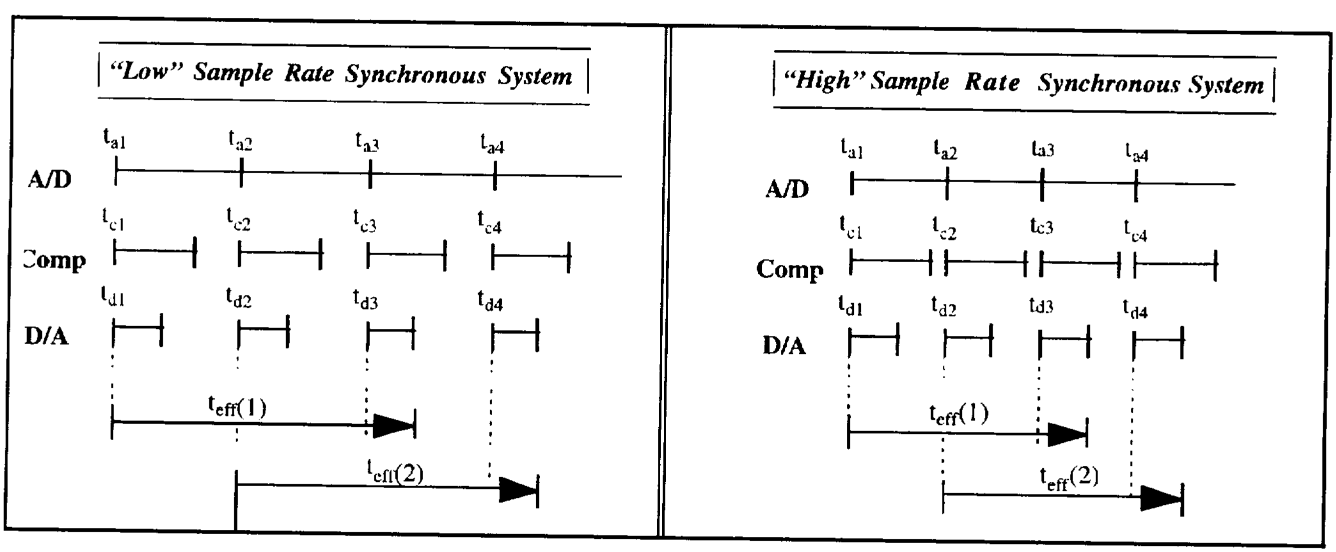


Figure 3, Synchronous System Event Timelines

The right side of Figure 3 shows that increasing the synchronous system sample rate will lower the effective latency. The effect of increasing that sample rate -- and decreasing the effective latency -- is summarized by example in Figure 4. The step response damping increases, the step response dominant resonant frequency increases, the frequency response gain (not shown) remains constant, and the frequency response phase at all frequencies increases; these effects can be summarized by stating that a linear decrease in the synchronous system effective latency results in a nearly linear increase in system phase margin. (These results are consistent with the well-documented discrete control system performance prediction algorithms.)

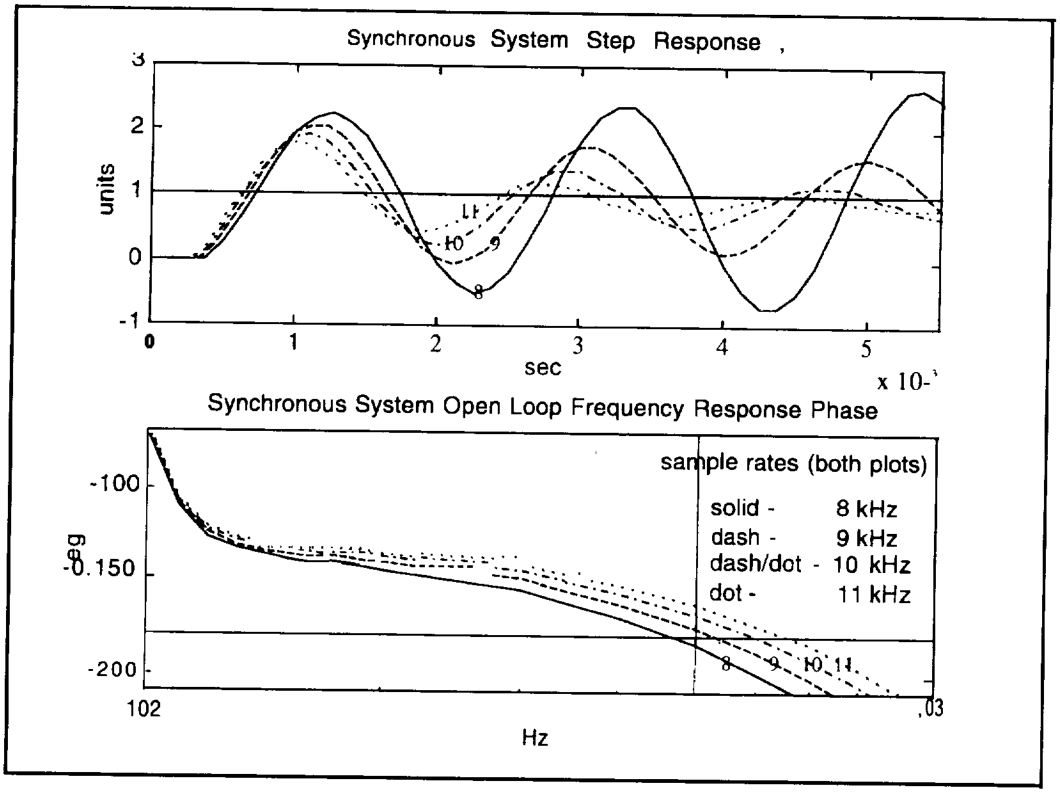


Figure 4. Synchronous System Step Response and Frequency Response Phase Plots

<u>Single Device Sample Rate Changes in a Multirate System</u>

Because the hardware and information update rate requirements for each of the individual components may vary, the control system designer may have the option to run each of those individual sampled-data devices at different rates [5]. By taking advantage of this option and releasing the individual sampled-data devices from their synchronous constraint, it is possible to greatly decrease the system effective latency. Unfortunately, this decrease comes with a penalty: the effective latency no longer remains constant from sample-to-sample. Figure 5 shows a cartoon example of the timing between individual component events in a multirate system; for simplicity, it is assumed from this point on that the time from initiation until completion of an individual event is negligible with respect to the sample time. The plot at the bottom of the figure demonstrates both the sample-to-sample variability and the overall periodicity [6] of the effective latency in an example multirate system.

If either the A/D or D/A sample rate inside a synchronous system is varied, the effective latency of the resultant multirate system also varies. Figure 6 presents the mean effective latency for two systems (one with the D/A and compensator sample rates set to 9 kHz, and the other with the D/A and compensator set to 11 kHz) versus a range of A/D sample rates. An interesting aspect of mean effective latency is its application as a simplistic measure of stability. Once the mean effective latency goes above a certain value (3.6045-4 seconds for the 9 kHz system, 3.7083-4 seconds for the 11 kHz system) the multirate system step response becomes unstable. If the individual component sample rates are adjusted to bring the mean effective latency below those values, the multirate system step response once again becomes stable.

The compensation/plant configuration, synchronous system performance metrics (phase margin, gain margin, and open-loop bandwidth), and individual component sample rates are all instrumental in calculating the maximum mean effective latency which will guarantee step response stability. To estimate that maximum mean latency for multirate systems configured as described in this paper, first predict the synchronous system phase margin.

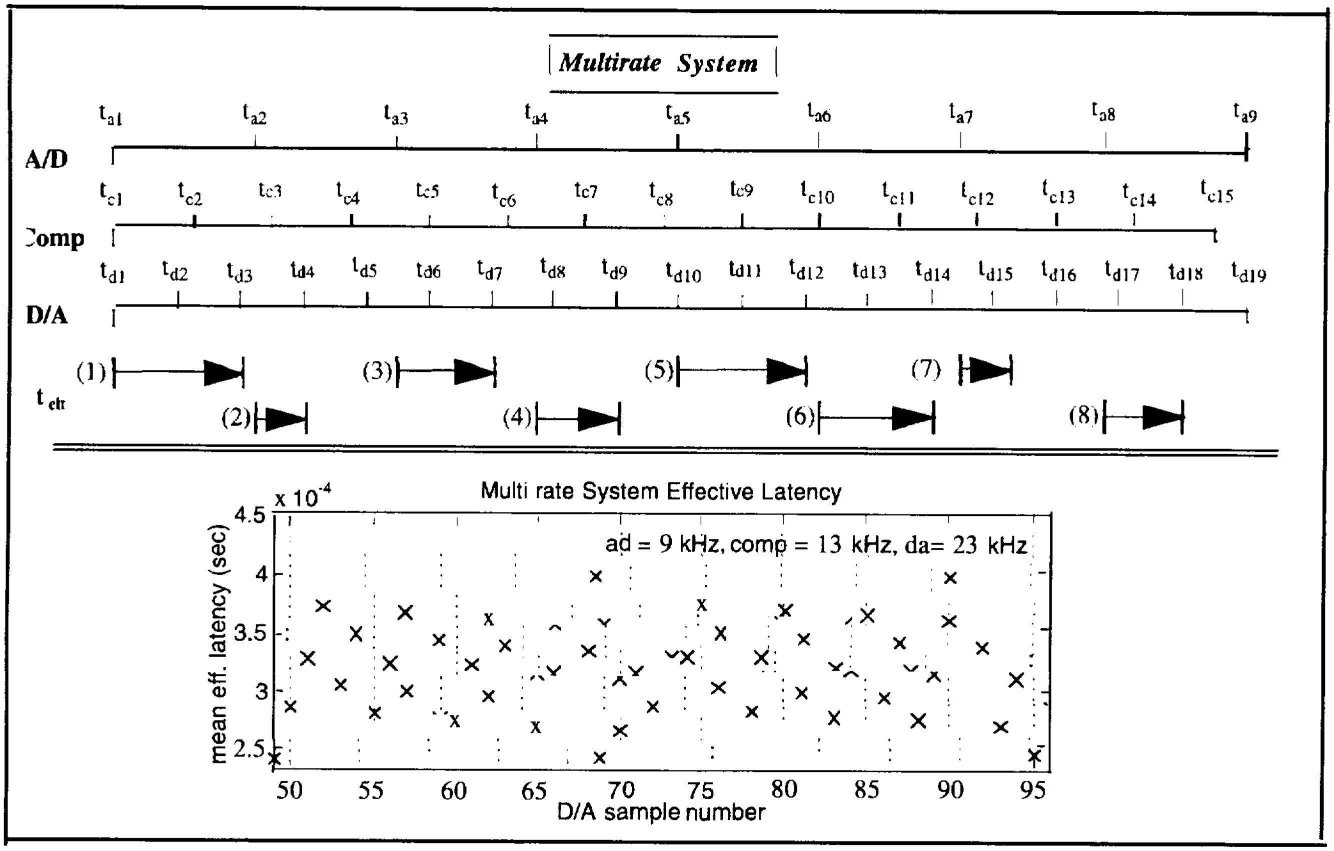


Figure 5. Multirate System Event Timelines

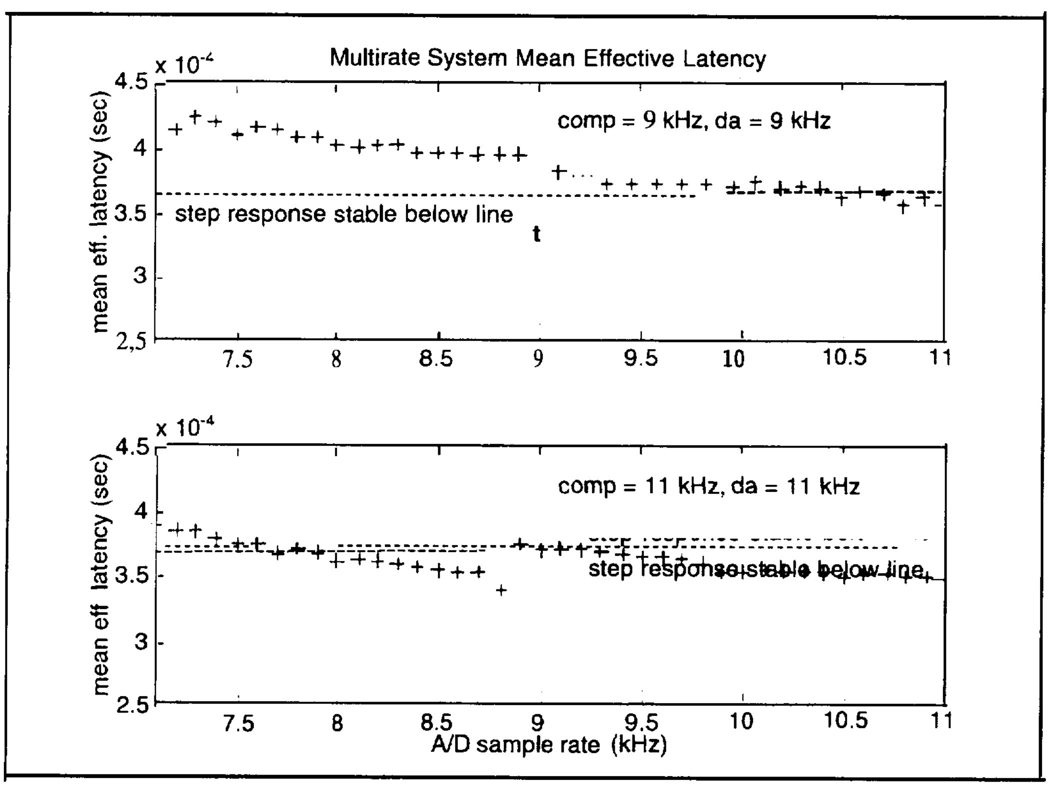


Figure 6. Multirate System Effective Latency (as a Function of A/D Sample Rate)

$$(pm)_{synch} = (pm)_{continuous} - [\,1260°(f_o/f_s)\,] \tag{1}$$

where pm = phase margin
f_o = continuous open-loop bandwidth
f_s = synchronous A/D, compensator, and D/A sample rates

(Note: The continuous system metrics used in equation (1) are calculated using an open-loop system which contains only s-domain representations of the plant and compensator.) Second, calculate the synchronous system mean effective latency.

$$(\,\overline{efflat}\,)_{synch} = [\,3/(f_s)\,] \tag{2}$$

where $\overline{efflat}$ = mean effective latency

Finally, estimate the multirate system maximum mean effective latency.

$$|\,(\,\overline{efflat}\,)_{multirate}\,|_{max} \;\; (\,\overline{efflat}\,)_{synch} + [\,(pm)_{synch}/(360° \times f_o)\,]\,| \tag{3}$$

Multirate systems with mean effective latencies lower than the minimum calculated above will generally produce stable step responses. The implication of this algorithm is that it is possible for the control system designer to predict the step response stability of a multirate system (given the restriction that two of the three component sample rates are equal) using the continuous system performance metrics and the multirate system mean effective latency.

Multiple Device Sample Rate Changes in a Multirate System

Once all the individual components are released from their synchronous constraint, the mean effective latency can be used to signal changes in system performance. As an example, a multirate system using the plant and compensator configuration from Figure 2 was built up; the A/D sample rate was set at 9.5 kHz, and the compensator sample was set at 12 kHz. Figure 7 shows the mean effective latency of that system with D/A sample rates ranging from 13 to 23 kHz. (The D/A sample rates are "normalized" by subtracting the compensator sample rate and then dividing the difference by the compensator sample rate.) It is clear that the mean effective latency is not a linear function of individual component sample rates; the mean latency is, however, subjected to distinct jumps at the major (i.e., halves, thirds, quarters) normalized sample rate fractions. (Note: there are similar jumps at smaller fractions, but those jumps are small enough to be ignored.)

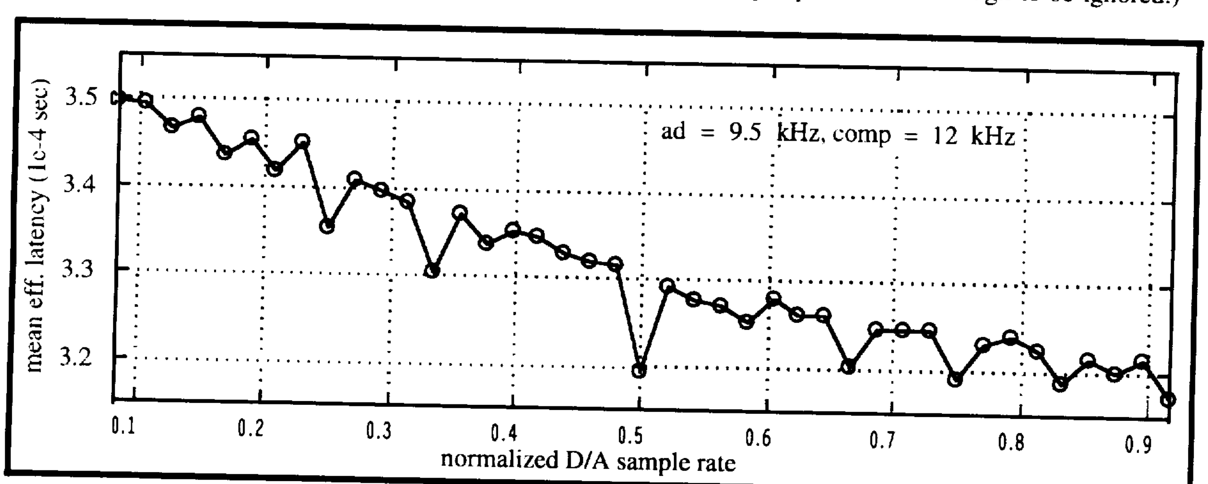


Figure7. Multirate System Mean Effective Latency

The distinct jumps shown in Figure 7 are also reflected in the system open-loop transfer function. As the normalized sample rate increases toward one-quarter, the frequency response only changes slightly; after it passes one-quarter, however, both the frequency response gain and phase change markedly to take on very different shapes. Continuing to increase the normalized sample rate does not markedly change the frequency response shape -- until the sample rate passes one-third. Figure 8 shows how the multirate system open-loop frequency response switches between two distinct curve shapes as the normalized D/A sample rate increases through major fractions. The plain and dashed lines in the figure are a result of normalized sample rates greater than 0.25, the lines with x's and o's have normalized sample rates greater than 0.33,

the line with +'s has a normalized rate greater than 0.5, and the line with *'s has a normalized rate greater than ().67.

For multirate systems configured as described in this paper, the following always hold true: the frequency response markedly changes from its original frequency response (the one calculated when the normalized sample rate equals O) to a perturbed shape once the normalized sample rate increases above 0.25, it returns to near the original response as the normalized sample rate increases above 0.33, it jumps back near the perturbed response as the rate increases above 0.5, it returns near the original response as the rate increases above 0.67, and moves back near the perturbed response above 0.75. The jumps in mean effective latency at major normalized sample rate ratios can, therefore, be used to signal the changes in multirate system behavior. The implication of these effective latency jumps is that it becomes simple for the control designer to tune the performance of' a multirate system; if the current normalized sample rate results in unacceptable system performance. slightly modifying that rate to the other side of a major fraction may improve performance to acceptable levels.

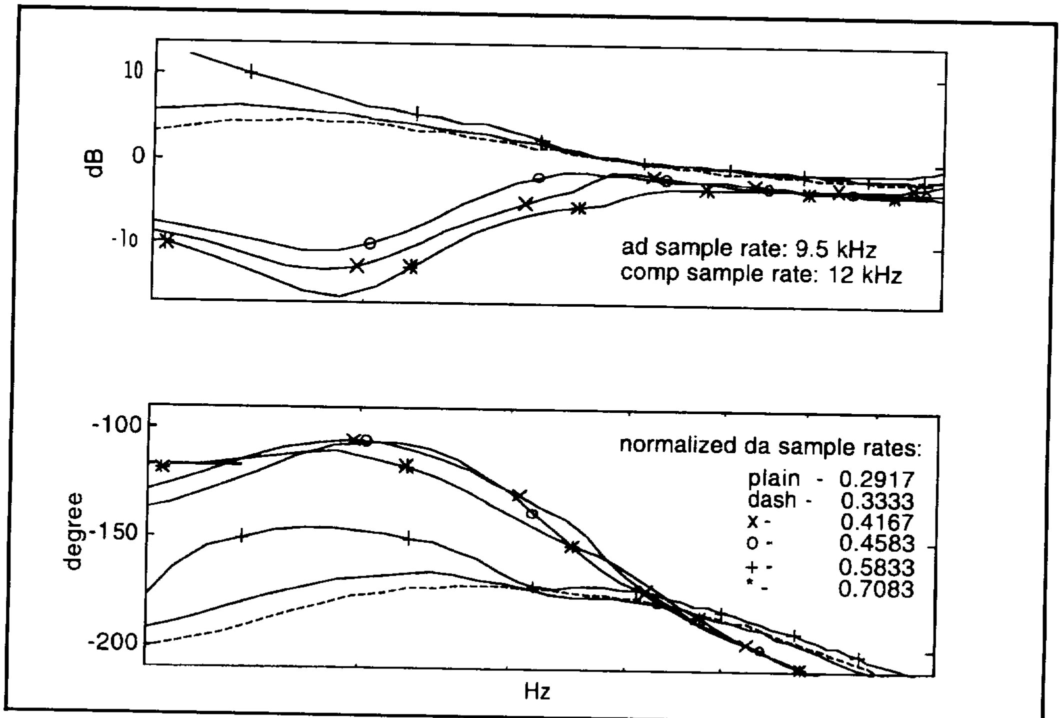


Figure 8. Multirate System Open-Loop Frequency Responses

Summary

The notion of effective latency -- the time it takes data to move from the sensor, through the A/D converter, compensator, and D/A converter, to the plant -- can be used to predict multirate system performance behavior. In multirate systems where only a single sampled-data device is run at a different sample rate than the other two devices, the system mean effective latency can be applied as a stability measure: if that mean latency rises above a certain value (dependent upon the plant, compensator, and individual component sample rates), the system step response will not be stable. In true multirate systems, where all sampled-data devices operate at different sample rates, the mean effective latency can be used to signal significant changes in multirate system performance behavior: as the normalized sample rate (D/A with respect to compensator) increases through major fractions (halves, thirds, quarters), the open-loop frequency response will actually switch between two distinct sets of gain and phase curves.

References

1. Godbout, Jr., L.F., Jordan, D., and Apostolakis, I.S, A Closed-Loop Model for General Multirate Digital Control Systems, IEEE Control, 1988, pages 494-499.
2. Berg, M. C., Mason, G. S., and Yang, G. S., A New Multirate Sampled-Data Control Law Structure and Synthesis Algorithm, Proceedings of the American Control Conference, 1991, pages 2749-2754.
3. Ling, K. V., Mashar, A., Lim, K. W., Experimental Evaluation of a Multirate Controller, Proceedings of the Asian Control Conference, Tokyo, 27-30 July 1994, pages 357-360.
4. Rekasius, Z.V., Digital Control with Computer Interruptions, American Control Conference, 1985, pages 1618-1621.
5. Wu, J, C., Chen, C. W., Rink, R. E., A New State-Space Representation of Multirate Sampled-Data Systems, American Control Conference, 1985, pages 1634-1639.
6. Ritchey, V. S., Stability of Asynchronous Multirate Linear Systems, Proceedings of the 27th Conference on Decision and Control, Austin Texas, December 1988, pages 1691-1696.

S0/33

LOW TEMPERATURE OPRATION OF A SWITCHING POWER CONVERTER

Carlos R. Anglada-Sanchez David Perez-Feliciano Biswajit Ray
University of Puerto Rico
Department of Electrical and Computer Engineering
Mayaguez, PR 00681-5000
Email: bray@sark.upr.clu.edu

Abstract

The low temperature operation of a 48 W, 50 kHz, 36/12 V pulse width modulated (PWM) buck dc-dc power converter designed with standard commercially available components and devices is reported. The efficiency of the converter increased from 85.6% at room temperature (300K) to 92.0% at liquid nitrogen temperature (77 K). The variation of power MOSFET, diode rectifier, and output filter inductor loss with temperature is discussed. Relevant current, voltage. and power waveforms are also included.

Introduction

With the recent development of high temperature superconducting materials with a critical temperature of approximately 125K, it is foreseeable that superconducting and low temperature electronics will find applications in future power transmission and motor-generator systems. Applications of low temperature electronics include cryogenic instrumentation, medical diagnostics, superconducting magnetic energy storage systems, and high speed computer, communication and electronic systems [1,2].

One particular application of interest to aerospace industries is in deep-space exploration programs. Typical passive spacecraft temperature varies from 448K in Mercury and gradually drops to 44K in Pluto. The current practice is to use radio-isotope thermoelectric generators (RTGs) and radio-isotope heating units (RHUs) to shield the electronic circuitry from low temperatures of deep space. In addition to the fact that RTGs and RHUs are expensive and environmentally unfriendly, RHUs are always on and therefore require additional heat rejection system when the space probe is near the earth orbit.

If electronic circuits can be designed to operate from room temperature down to a very low temperature (e.g., liquid nitrogen temperature of 771<), then the use of RTGs and RHUs will not be required. Additionally, launch weight and cost wil I be significantly reduced due to the absence of any heat rejection and thermal shielding systems for electronics.

The objective of this research was to investigate the possibility of designing and operating basic power processing electronics suitable for operation from 3001< down to 77K using standard commercially available components. For this purpose, a 48 W, 36±6/12 V, 50 kHz pulse-width modulated (PWM) dc-dc power converter was designed and tested in the laboratory at 3001< and 771<. Its efficiency increased from 85.6% at 300K to 92% at 77K. The converter performance, component loss, and relevant waveforms are discussed in the following sections.

PWM Buck converter

The 48 W, 36±6/12 V, 50 kHz PWM buck converter was designed for operation at 300K as well as at 77K. This type of power converter can potentially be used in small scientific/experimental spacecraft such as the proposed

* This work was supported by NASA through the Tropical Center for Earth and Space Studies, a University Research Center at the University of Puerto Rico - Mayaguez Campus.

CLIR (Combined Lander and Instrumented Rover). The basic converter circuit is shown in Fig. 1. It was designed for a minimum load current of 1 A for continuous conduction mode of operation and a maximum output voltage peak-to-peak ripple of 1%.

Based on the steady-state analysis for continuous conduction mode of operation of a PWM buck converter [3], the following design equations are used for the power circuit design:

$$L_f \geq \frac{V_o(1 - D_{\min})}{2 f_s I_{o,\min}} \tag{1}$$

$$C_f \geq \frac{(1 - D_{\min}) V_o}{8 L_f f_s^2 (\Delta V_o)} \tag{2}$$

where, D_{min} = minimum duty ratio $\equiv V_O/V_{in,max}$, $I_{O\,min}$ = minimum output (load) current for continuous conduction mode of operation, f_s = switching frequency, and ΔV_O = peak-to-peak output ripple voltage.

Based on equation (1), the required filter inductance is 85.7 μH and was designed using a distributed air gap ferrous alloy powder core from Magnetics (Kool Mμ 77934) with a relative permeability of 90. Standard magnet wire (3*#20 AWG) was used for winding 27 turns on the core to obtain the desired filter inductance. Based on previously reported work with molypermalloy powder cores [4], it was expected that the designed inductor will work at low temperature even though its loss might increase a little [5,6].

The required filter capacitance was found to be 41.7 μF using equation (2). For the power converter an output filter capacitance of 50 μF was used. Standard metallized polypropylene film capacitors were used because of their superior low temperature characteristics [7].

Power semiconductor selection: For low temperatures, the primary semiconductor material is silicon, although gallium-arsenide also has considerable potential and the primary device is the field-effect transistor in various forms [1,2,8]. Reduced temperature operation offers improvements in performance through improvement of material-based properties such as electronic carrier mobility, thermal conductivity, and electrical conductivity. For this work, standard plastic packaged (TO-220) IRF-540 power MOSFET (28 A, 100 V, 85 mΩ, 560 pF device) was used as the controllable power switch and a MBR-20100CT (20 A, 100 V, V_F = 0.9 V) Schottky diode was used as **the** output rectifier.

Experimental procedure and results:

The complete power converter tested at room and liquid nitrogen temperature is shown in Fig. 2. The open-loop control circuit was designed around a bicmos voltage-mode PWM IC (TC35C25CPE) and a cmos driver IC (IR-2113). The programmed switching frequency was 50 kHz, and the duty ratio was controllable through the 20 kΩ potentiometer kept at room temperature. The driver IC has independent high and low-side referenced output channels. The high-side floating channel was used to drive the power MOSFET without having to use an opto-coupler or an isolation transformer. The resistors used in the control circuit were metal-film type and the capacitors were NPO ceramic type, both having fairly temperature independent characteristics [7].

The full-load data were recorded both at 300K and 77K. The control and power circuitry were placed inside a Dewar flask whereas the measuring and sensing instruments were at room temperature, resulting in a highly non-compact circuit layout. The power converter was capable of restarting at 77K. Recorded results and waveforms are discussed next.

The full-load efficiency of the converter increased from 85.6% at 300K to 92% at 77K, and the converter loss decreased from 6.9 W at 300K to 3.6 W at 77K. The switch loss was less than a watt at both temperatures. The filter inductor loss was about two watts and increased slightly at 77K compared to 300K, primarily due to increased

flux density and decreased core resistivity [5,6]. The rest of the loss was due to the output rectifier, output filter capacitor, and stray losses in long cables used for testing the circuit immersed in liquid nitrogen. The stray loss decreased significantly at 77K compared to 300K, primarily due to improved electrical conductivity of wires.

Figures 3, 4, and 5 show the voltage, current. and power waveforms of switch, diode, and inductor, respectively, for 77K and 3001< operation. The voltage waveforms show a significant amount of high frequency ringing during the switch turn-on and turn-off instants. This is inherent in any hard-switched converter because instantaneous change in voltage and current is opposed by the switch capacitance and circuit layout inductance, respectively, The high frequency ringing is caused by resonance between the switch and Schottky rectifier capacitance and the circuit layout inductance [5]. The circuit layout inductance was significant due to the experimental setup where the converter circuit was immersed in a liquid nitrogen filled Dewar whereas all the measuring instruments were outside the Dewar. This is evident in Figs. 3 and 4, because the turn-off ringing is much more significant than the turn-on ringing. It can also be noticed that the waveforms at 77K and 3001< look almost identical except for the fact that the switching noise is somewhat less at 77K.

Finally, the switching frequency of the converter changed from 51.3 kHz at 300K to 49,7 kHz at 77K, indicating an overall change of only 3°/0. This change is primarily contributed by the external timing resistor and capacitor used for programming the switching frequency of the PWM control IC. The monolithic resistors and capacitors internal to the PWM IC also play an important role. However, a minor variation in switching frequency will not be a problem for most switching power converters.

Conclusions

This work demonstrated the possibility of successfully operating a power electronic converter from room temperature down to the liquid nitrogen temperature, designed with standard commercially available components and devices. 'The efficiency of the converter increased from 86.6% at 300K to 92% at 77K. The switching frequency decreased by 3% at 77K respect to 300K operation. The semiconductor and passive component loss as a function of temperature must be looked at further details. The electronic packaging and related thermal mismatch issues must be investigated before any low temperature use of electronics is contemplated.

References

1. Deen, M. J., "Low temperature microelectronics: opportunities and challenges," Proc. Symp. Low Temperature Electronic Device Operation, The Electrochemical Society, Vol. 91-14, pp. 25-37, 1991.
2. R. K. Kirschman, "Low-temperature semiconductor electronics," Proc. Symp. Low Temperature Device Operation, The Electrochemical Society, Vol. 91-14, pp. 1-12, 1991.
3. N. Mohan, T. M. Undeland, and W. P. Robbins, "Power Electronics: converters, applications, and design," 2nd edition, John Wiley, pp. 164-172, 1995.
4. Ray, B., et. al., "Low temperature performance of a boost converter with MPP and HTS inductors," IEEE Applied Power Electronics Conference Record, pp. 883-888.1996.
5. F. W. Ackermann et. al., "Magnetic properties of commercial soft magnetic alloys at cryogenic temperatures," Adv. Cryogenic Engineering, Vol. 16, pp. 46-50, Plenum Press, 1971.
6. J. J. Gniewek and R. L. Powell, 'Temperature dependence of magnetic losses," Adv. Cryogenic Engineering, Vol. 7, pp. 303-310, Plenum Press, 1962.
7. Blanchard, R. and Severns, R., "Designing switch-mode power converters for very low temperature operation," Proc. Powercon, Vol. 10, pp. D2.2-D2.11, 1983.
8. Muller, O. "On-resistance, thermal resistance, and reverse recovery lime of power MOSFETs at 771<," Cryogenics, Vol. 29, pp. 1006-1014, 1989.

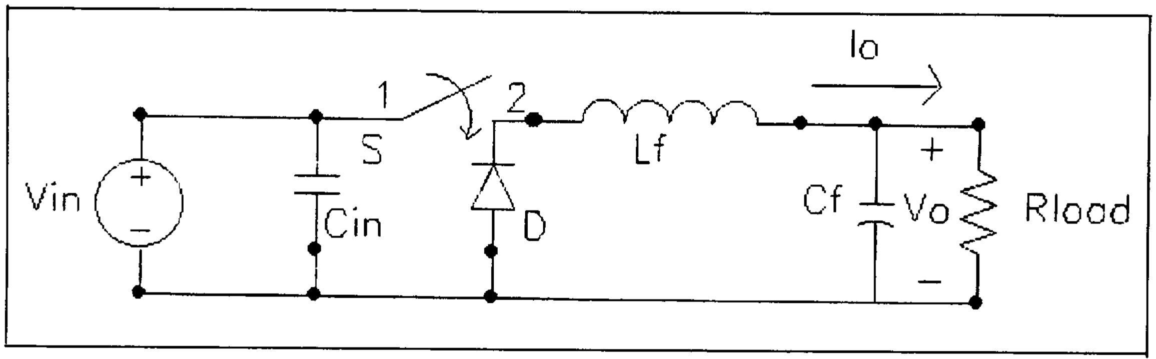


Figure 1 PWM buck dc-dc converter.

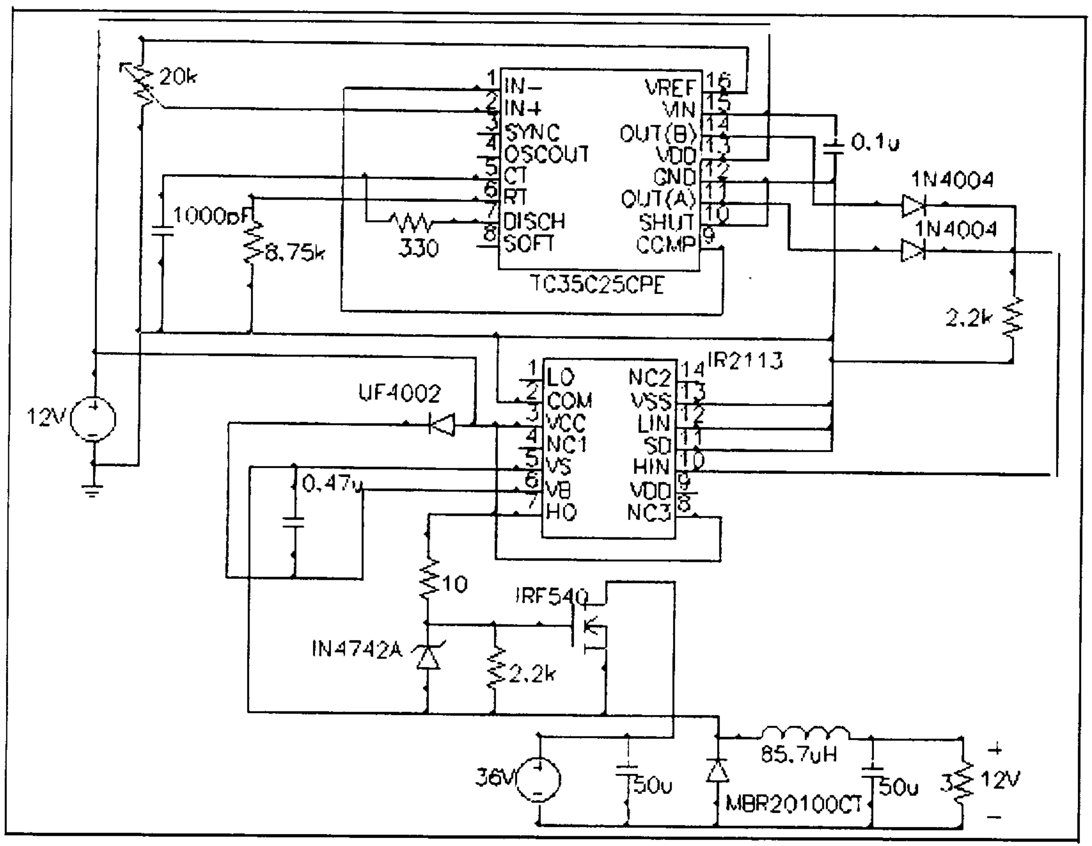
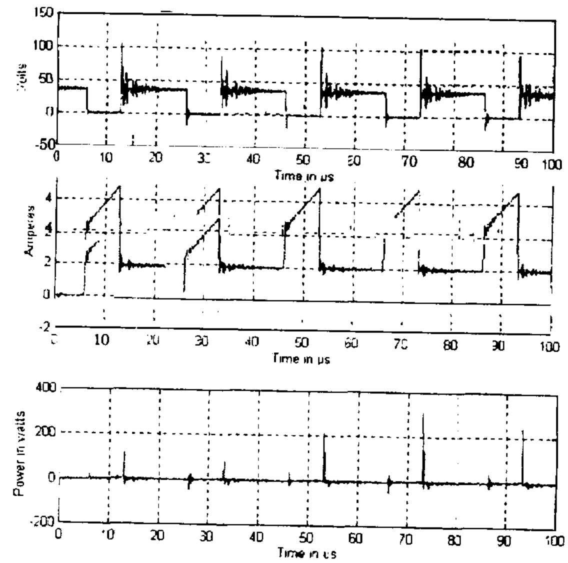
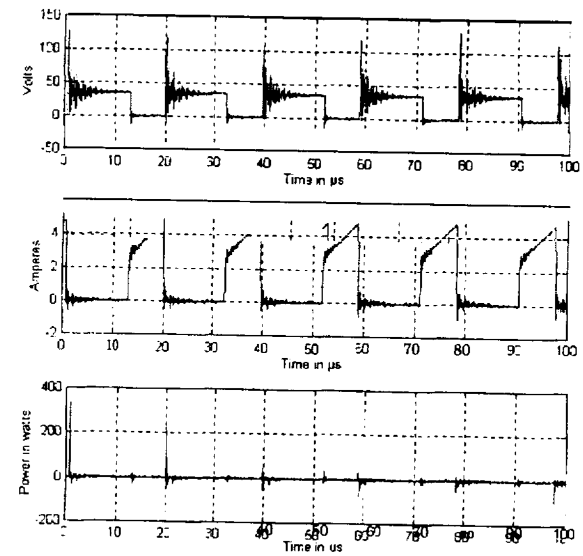


Figure 2 Complete converter circuit tested at 300K and 77K.

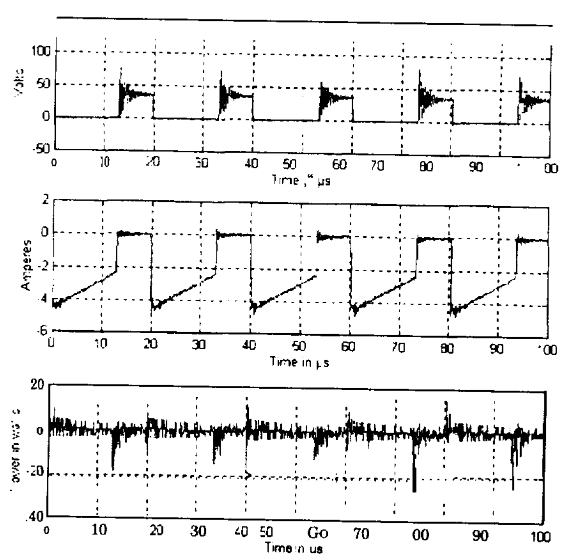


(a) 77K operation

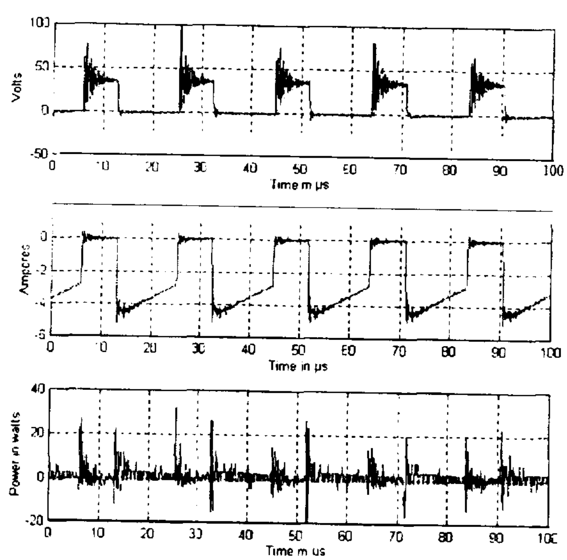


(b) 300K operation

Fig. 3 Switch voltage, current, and power waveforms.

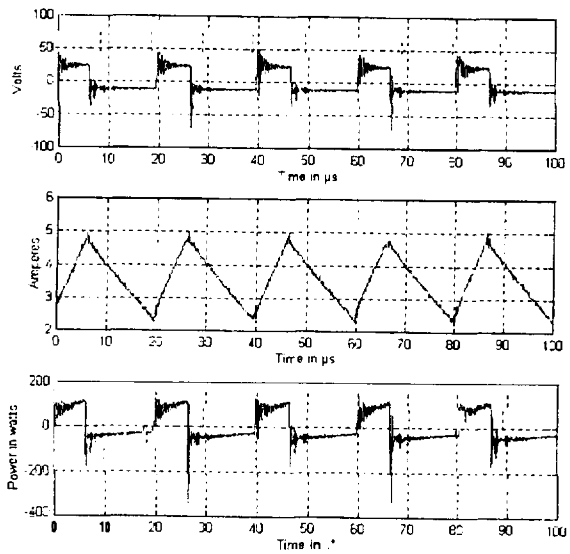


(a) 77K operation

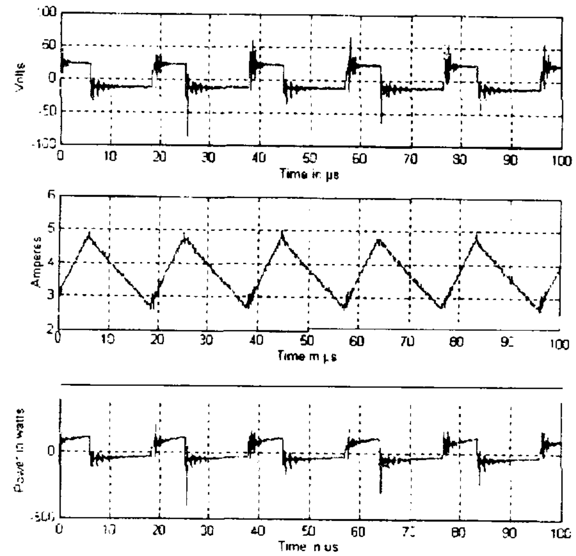


(b) 300K operation

Fig.4 Diode voltage, current, and power waveforms.



(a) 77K operation



(b) 300K operation

Fig. 5 Inductor voltage, current, and power waveforms.

Sensor Fault Detection and Accommodation Using Causal Probabilistic Networks for Autonomous Control Systems

Hrishikesh Aradhye

A. Sharif Heger

Department of Chemical and Nuclear Engineering

Center for Autonomous Control Engineering

The University of New Mexico

Abstract

This paper discusses the result of our research on the use of Bayesian Belief Networks (BBN) for sensor Fault Detection and Accommodation (FDA). This research supports one of NASA's areas of concern and active research for autonomous spacecraft guidance and control. Sensor fault detection and accommodation aims at using hardware and/or analytical redundancies for detecting sensor failures and integrating available redundant information for improving the reliability of the data that goes to the input of automatic control lers or human operators. Bayesian Belief Networks are directed acyclic graphs with the dependencies and independencies among the variables imbedded within the network structure. They simulate the fl ow of human reasoning through propagation of evidence via sound probabilistic mathematics. Using BBNs, we have developed several FDA systems ranging from static, non-learning models of discrete variables to dynamic, learning models applicable to continuous variables. We present a review of literature of current research efforts in this area and discuss the results of our method.

Introduction

For space missions and the emergent space commercialization, NASA and the space industry need autonomous learning control systems that can optimize the performance of the spacecraft subject to the constraints of their missions, environmental conditions, and instrument input. To this end, we have launched a research project to develop fault detection and accommodation techniques using Bayesian Belief Networks. These FDA systems have evolved from static, non-learning models of discrete variables to dynamic, learning models applicable to continuous variables. These models can be used for autonomous control applications in areas such as safety-critical fault detection and isolation, signal validation advisor augmenting aircraft flight data recorder, or onboard engine status advisor.

Accurate information about the state of a system is important for optimum tracking or regulatory control of the system, as well as safety considerations. Sensor errors or failures can lead to wrong control actions, or may mask a process malfunction. A fault is defined as "a nonpermitted deviation of a characteristic property which lead to the inability to fulfill the intended purpose" [Isermann, 84]. Therefore, recognizing a fault and, consequently, reorganizing the system to mitigate the effects of the fault is the objective of Fault Detection and Accommodation. Even in absence of a fault, the data reconciliation module of FDA system can provide a closer estimate of the variable under considerate ion, thus leading to better control. This is important because of inherent error associated with the sensor output even during its normal operation.

The two approaches used for FDA are based on hardware and/or analytical redundancies. Hardware redundancy approach suggests the use of multiple sensors for the same function, and obtains an estimate by data-fusion of these sensor readings. The analytical or functional redundancy approach assumes or learns a functional relationship, i.e., a mathematical model for the system variables. A comparison between the actual sensor readings and model predictions leads to the basis for failure detection. In addition to traditional methods [Sturza, 90; Chin, 93; Anderson, 86; Milner, 86; Barker, 91; Litt, 94; Levesque, 93] and statistical data fusion techniques [Park, 93], the emerging paradigms of fault detection are AI-based. These can be roughly categorized as the following:

1. **Knowledge Based Systems-These** expert systems use symbolic reasoning to encode an expert's/operator's knowledge into rules, with certain uncertainty factors associated. They are specific to the

process at hand, and require extensive experience on the part of the operator [Watton, 94; Lee, 94; Russell, 89; Singer, 93; Quinn, 91; Hudlicka, 84; Allen, 90; JPL, 94; Collins, 90] .

2. **Artificial Neural Networks**—ANNs **have been** proven to function as efficient classifiers, optimizers, and function approximators. Several learning algorithms and network structures have been proposed, and more are being worked on in the present time. Fault detection has been presented as a classification problem, a clustering problem, or a function approximation one, leading to different network architectures. [Bishop, 91; Mourot, 93; Chiang, 93; Chen, 94; Napolitano, 95; Hoskins, 9 1; Karjala, 94; Karjala, 92; Karjala, 96; Watanabe, 94; Abdel Mageed, 93; Bernieri 94]

3. **Fuzzy Logic—The** speed and simplicity of approximate reasoning and calculation has been the basis of the fuzzy logic domain. Solutions to fault detection problem with this approach have been proposed as application of fuzzy rule bases on residual readings, or along with neural nets using the functional redundancy approach [Dexter, 93; Singer, 91; Mourot, 93; Goode, 93; Sauter, 94; Holbert, 95; Heger, 96]

4. **Bayesian** Belief **Networks**—This approach is relatively new. Its use in the past has been limited to discrete domain, where the network structure encodes the rules defining the system. These networks can also be dynamic, leading to on-line and localized learning [Deng, 93; Kirch, 94; Nicholson, 94; Nicholson, 96]

Bayesian Belief Networks

Bayesian Belief Networks (also known as Causal Networks and Probabilistic Networks) are acyclic directed graphs (DAG). Each node in the network represents a continuous or discrete propositional variable. The link between two nodes denotes a causal dependency of the child on the parent. The network functions through a causal reasoning by propagation of "diagnostic" or "causal" evidence from the variables with known values. Given the evidence, the belief in a variable taking its values is computed as a probability distribution.

We have chosen BBNs to model FDA systems for the following reasons:

1. The nodes in a causal network are "meaningful" as they represent actual events or processes. This is in contrast to the nodes in a neural network where they could have no physical meaning. The connections reflect the "cause-effect" relationships among the nodes, and the propagation of evidence through probabilities is similar to human flow of reasoning. In that respect, the function of causal networks is more explainable, understandable, and acceptable to a human operator.

2. once the network is constructed, its operation is based upon sound theorems of probability. This eliminates the requirement of intuition and insight associated with quasi-probabilistic measures such as rule-weights or confidence factors, used in fuzzy or traditional expert systems.

3. The dependencies and independencies are implicit in the structure of the causal network. Addition or removal of an instrument can be taken care of by introduction and removal of a node. along with its conditional probability table. It does not affect the structure of the rest of the network, and does not require any changes in any other nodes, connections, or tables. On the contrary, upon introduction of a change, a neural network will have to be retrained, and a fuzzy or traditional rule base will have to eliminate or introduce the calculation of residuals for the sensor in question.

4. Time intervals between successive sensor measurements may be different for different sensors. Readings for not all the sensors may be available at any given instant. Causal networks implicitly allow for missing data, while such an incorporation has to be specifically introduced as rules in a rule-based system, and is very difficult for neural networks.

Despite its benefits, there are some limitations associated with the use of BBNs. For example:

1. All variables are assumed to be random, with precisely or imprecisely known distributions. Some of the variables may not be random, and may be varying as a deterministic function of time.

2. We assume the availability of an underlying model, mathematical or otherwise. This is in contrast to the approach of neural works where no model is assumed and is learned from data. Given the initial model, the BBN can improve it based on its experience.

3. The existing- theory allows the use variables with continuous values subject to two restrictions: the probability distributions of these variables must be Gaussian, and the functional dependencies among the nodes must be linear. Thus, any non-linear relationship must be locally linearized.

4. No loops are allowed in the network, i.e., the graph must be a DAG.

Example

To illustrate the effectiveness and simplicity of BBNs in FDA, consider a simple system in which an electric heater is used for heating a fluid to a specified temperature, as shown in Figure 1. All variables except the mass flow rate of the inlet fluid and the current in the heating coil are assumed to be constant. Further, all time delays are neglected. We assume the electric current and coolant mass flow rate are measured by an ammeter (AM) and a calibrated rotameter (RT), respectively, Three sensors are used to measure the outlet temperature (THM 1, THM2 and THM3). A uniform distribution for all *a priori* probabilities is assumed. The goal is the dynamic detection and accommodation of the failures of the four sensors, using the available data.

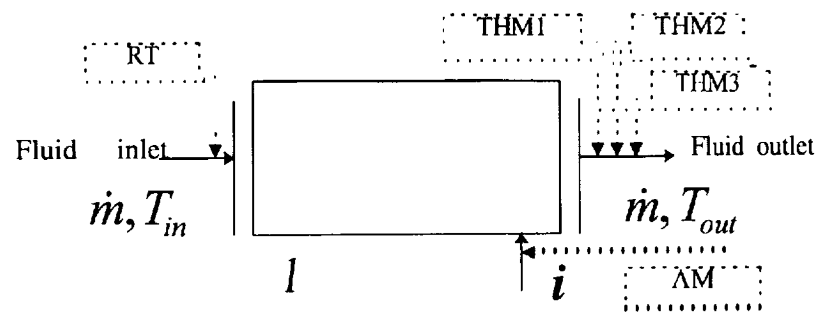


Figure 1: A schematic diagram of the example heating system.

The steady state equation characterizing the system is:

$$T_{out} = T_{in} + \frac{i^2 R}{\dot{m} C_p} .$$

Small perturbations in current and mass flow rate, around the steady state, yield a linearized expression for the resultant change in the outlet temperature. When dedimensionalized, it gives:

$$\theta = 2\eta - \mu ,$$

where

$$\theta = \frac{T_{out} - \overline{T}_{out}}{\overline{T}_{out} - T_{in}}, \eta = \frac{i - \bar{i}}{\bar{i}}, \text{ and } \mu = \frac{\dot{m} - \overline{\dot{m}}}{\overline{\dot{m}}} .$$

$\overline{T}_{out}$, $\bar{i}$ and $\overline{\dot{m}}$ be the mean values of T_{out}, i, and m, respectively. In the simulations, measurement of each variable is subjected to normally-distributed small sensor noise. The sensors are made to fail with fixed probabilities of failure. Simultaneous failures are allowed to occur. The error of a failed sensor has a standard deviation equal to 3.16 times that of a sensor in its normal mode. Upon failure, a sensor remains in the failed mode until the failure is detected, and then returns to its normal mode. A failure detection alarm is raised when the belief about the sensor being in the failed state rises above a threshold of 0.9.

The BBN representation of this FDA problem is shown Figure 2. The partial result of the analysis is shown in Figure 3. For a 2000-step experiment, the five sensors experienced 214 failures which were detected by the FDA system. In addition, there were nine false alarms, i.e., the FDA system sounded the alarm when the particular sensor was actually in operating status.

Current Results

We have developed a dynamic, learning BBN method for both steady and unsteady state failure detection and accommodation. Hugin API [HUGIN, 95] has been used for this purpose for efficient compatibility between a C++ simulator and the network. The following variations of the system definition have been investigated:

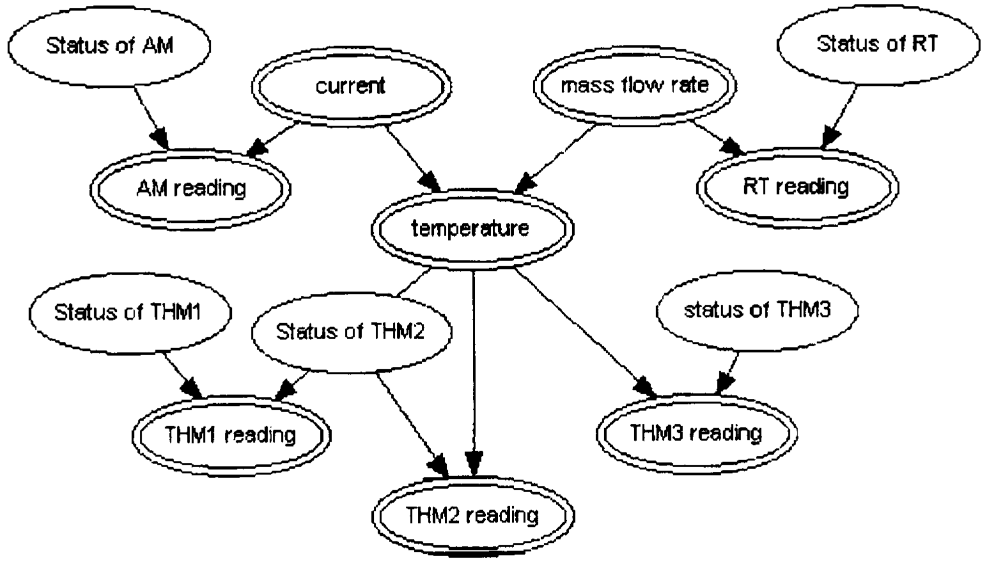


Figure 2: BBN representation of the electrical heater system.

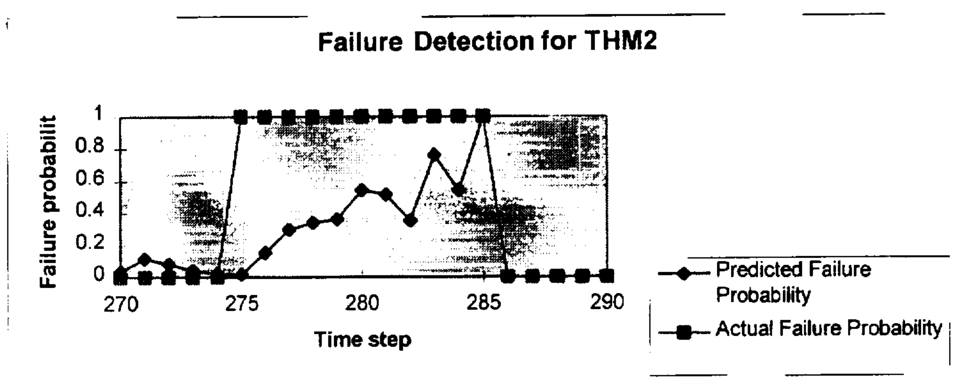


Figure 3: A comparison between the actual state of sensor TMH2 and the corresponding belief of the FDA system. As the sensor stays in the failed state, the belief of the system continues to grow until it reaches the threshold value when it sounds an alarm,

1. discrete-valued variables—single variable, single sensor problem; single variable, multiple sensors problem (hardware redundancy); multiple variables, one sensor per variable problem (functional redundancy); and multiple variables, multiple sensors problem (hardware and functional redundancy)
2. continuous-valued variables—multiple continuous variables with linear dependencies; and multiple continuous variables with nonlinear dependencies
3. learning and dynamic causal networks
4. causal fault detect ion simulations—steady state, three variables problem: the electric heater system; and dynamic simulation of the electric heater system fault detection
5. process control and causal Fault Detection and Accommodation-outlet temperature regulatory control for the electric heater system.

Our future plans call for testing our system using an appropriate NASA flight control problem. We believe that our results would be beneficial in many industrial applications, specifically those of NASA in its space missions. For space missions and the emergent space commercialization, NASA and the space industry needs autonomous learning control systems that can optimize the performance of the spacecraft subject to the constraints of their missions, environmental conditions, and instrument input. Our research can lead to the development of an in-flight

fault detection, stabilization, and reconfigurable control that will monitor safety-critical signals to support better diagnostics and pinpoint emerging problems to improve spacecraft performance for manned and unmanned missions. In addition to meeting the needs of NASA in this important area, the developed product will have commercial applications in aerospace, power generation, chemical and petrochemical processing, and automotive industries.

References

M. F. Abdel Mageed, A. F. Sakr, and A. Bahgat, "Fault Detection and Identification Using a Hierarchical Neural Network-based System", in the Proceedings of the IECON'93: International Conference on Industrial Electronics, Control and Instrumentation, 1993, vol. 1, pp. 338-343.

S.M. Allen. "Expert System Approach to Global Fault Detection and Isolation Design", Report Number: WL-TR-91-3007, 1990.

W.E. Anderson, "Technical Contributions to the Development of Incipient Fault Detection/Location Instrumentation", Report Number NBSIR-86/3392, 1986.

R. W. Barker, "Incipient Fault Detection Using Higher-Order Statistics", Doctoral thesis, AFIT/CI/CIA-91-023D, 1991.

A. Bernieri, G Betta, A Pictrosanto, and C. Sansone, "A Neural Network Approach to Instrument Fault Detection and Isolation", in Proceedings of the 1994 IEEE Instrumentation and Measurement Conference, 1994, pp. 139-144.

T. Bishop, "Space Shuttle Main Engine Fault Detection Using Neural Networks", Report Number: N9 1-20816, 1991.

J. Chen, R.J. Patton, and G. P. Liu, "Detecting Incipient Sensor Faults in Flight Control Systems", in the Proceedings of the 1994 IEEE Conference on Control Applications, 1994, vol. 2, pp. 871-876.

C. Chiang, J. Juang, and H. Youssef, "Neural fault Detection, Isolation, and Estimation Design", in the Proceedings of the 1993 International Joint Conference on Neural Networks, 1993, vol. 2.

H. Chin, "Fault Detection of Helicopter Gearboxes Using the Multi-Valued Influence Matrix Method", NASA, Report Number NASA-TM- 106100, 1993.

G. Collins, "Research on Intelligent Fault Detection and Repair Advisory Systems", Report Number: AFOSR-TR-91-0499, 1990.

R. H. Deng, A. A. Lazar, and W. Wang, "A Probabilistic Approach to Fault Diagnosis in Linear Lightwave Networks", IEEE Journal on Selected Areas in Communications, vol. 11, no. 9, pp. 1438-1448, 1993.

A. L. Dexter, "Fault Detection in HVAC Systems Using Fuzzy Models", Report Number: AN25/UK/15 1092/1.1, 1993.

P. V. Goode, and M. Chow, "Neural/Fuzzy Systems for Incipient Fault Detection in induction Motors", in the Proceedings of the IECON'93: International Conference on Industrial Economics, Control and Instrumentation, 1993, vol. 1, pp. 332-337.

A. S. Heger, K.E. Holbert, and A. M. Ishaque, "Fuzzy Associative Memories for Instrument Fault Detection", *Annals of Nuclear Energy, vol. 23,* no. **9**, pp. *739-756, 1996.*

K. E. Holbert, A. S. Heger, and A. M. Ishaque, "Fuzzy Logic for Power Plant Signal Validation", in the Proceedings of the Power Plant Dynamics, Control and Testing Symposium, 1995, pp. 20.01.

J. C. Hoskins, K. M. Kaliyur, and D. M. Himmelblau, "Fault Diagnosis in Complex Chemical Plants Using Artificial Neural Networks", *AIChE Journal,* vol. *37,* no. 1, pp. 137-141, 1991.

E. Hudlicka, Design of a Knowledge-Based Fault Detection and Diagnosis System", Report Number: TR-84-03, 1984.

HUGIN API Reference Manual, Version 2.0, 1995.

R. Isermann, "Process Fault Detection Based on Modeling and Estimation Methods-A Survey", *Automatic, vol. 20,* no. *4,* pp. 387-404, 1984.

Jet Propulsion Laboratory, "Hidden Markov Models and Neural Networks for Fault Detection in Dynamic Systems", Report Number: N95-25261, 1994.

T. W. Karjala, D. M. Himmelblau, and R. Miikkulainen, "Data Rectification Using Recurrent (Elman) Neural Networks", in the Proceedings of the IEEE-IJCNN, vol. 2, Baltimore, pp. 901-906, 1992.

T. W. Karjala, and D. M. Himmelblau, "Dynamic Data Rectification Using the Extended Kalman Filter and Recurrent Neural Networks", in the Proceedings of the IEEE-ICNN, vol. 5, Orlando, pp. *3244-3249,* 1994.

T. W. Karjala, and D. M. Himmelblau, "Dynamic Rectification of Data via Recurrent Neural Nets and the Extended Kalman Filters", *AIChE Journal,* vol. *42,* no. **8**, pp. 2225-2239, 1996.

H. Kirch and K. Kroschel, "Applying Bayesian Networks to Fault Diagnosis", in the Proceedings of the IEEE Conference on Control Applications, 1994, vol. 2, pp. 895-900.

S. C. Lee, "Sensor value Validation Based on Systematic Exploration of the Sensor Redundancy for Fault Diagnosis KBS", *IEEE Transactions on Systems, Man, and Cybernetics, vol. 24,* no. 4, pp. 594-605.1994.

M.J. Levesque, "Fault Detection and Isolation for the Bluebird Test Bed Aircraft", Master's Thesis, Report Number: AD-A277 979, 1993.

J. Litt, "Sensor Fault Detection and Diagnosis Simulation of a Helicopter Engine in an Intelligent Control Framework, Report Number NASA-TM- 106784, 1994.

G.M. Milner, "Incipient Fault Detection Study for Advanced Spacecraft Systems", Report Number NASA-CR-172014, 1986.

G. Mourot, S. Bousghiri, and F. Kratz, "Sensor Fault Detection Using Fuzzy Logic and Neural Networks" in the Proceedings of the 1993 IEEE International Conference on Systems, Man and Cybernetics, 1993, pp. 369-374.

M. R. Napolitano, C. Neppach, V. Casdorph, and S. Naylor, "Neural-Network-Based Scheme for Sensor Failure Detection, Identification, and Accommodation", *Journal of Guidance, Control and Dynamics,* vol. 18, no. 6, pp. 1280-1286, 1995.

A. E. Nicholson, and J. M. Brady, "Dynamic Belief Networks for Discrete Monitoring", *IEEE transactions on systems, man, and cybernetics, vol. 24,* no. 11, pp. 1593, 1994.

A. E. Nicholson, "Fall Diagnosis Using Dynamic Belief Networks", in Lecture Notes in Computer Science, vol. 1114, pp. 206, 1996.

S. Park, and C. S. George Lee, "Fusion-based Sensor Fault Detection", in the Proceedings of the 1993 IEEE International Symposium on Intelligent Control, 1993, pp. 156-161.

T.M. Quinn, "Autonomous Fault Detection, Isolation, and Recovery System for a 20-Khz Electric Power Distribution Test Bed", Report Number: NASA-TM-104344, 1991.

B.D. Russell, "Advanced Power System Protection and Incipient Fault)etection and Protection of Spaceborne Power Systems", Final Report: NASA-CR-1 85330, 1989.

D. Sauter, N. Mary, F. Sirou, and A. Thieltgen, "Fault Diagnosis ir Systems Using Fuzzy Logic", in the Proceedings of the 1994 IEEE Conference on Control Applications, 1994, vol. 2, pp. 883-888.

R. M. Singer, "Pumping system fault detection and diagnosis utilizing pattern recognition and fuzzy inference techniques", Report Numbers: ANL/CP-7 1255; 1991.

R.M. Singer, "Applications of pattern recognition techniques to online fault detection", Report Number: ANL/RA/CP-79576, 1993.

M.A. Sturza, "Fault Detection and Isolation (FDI) Techniques for Guidance and Control Systems", Report Number: N90-29366, 1990.

K. Watanabe, S, Hirota, L. Hou, and D. M. Himmelblau, "Diagnosis of Multiple Simultaneous Fault via Hierarchical Artificial Neural Networks", *AIChE Journal,* vol. 40, no. 5, pp. 839-848, 1994.

J. Watton, O. Lucca-Negro, and J. C. Stewart, "An On-line Approach to Fault Diagnosis of Fluid Power Cylinder Drive Systems", *Journal of Systems and Control Engineering, vol. 208,* pp. 249-262, 1994.

URC97012

Detection of Human Subject of Interest in a Digital Image via Fuzzy Logic and Color Segmentation

Ali Asgharzadeh[1], Ali Jadbabaie, Babak Firoozbakhsh, and Eddie A Bononcini[2]

NASA Center for Autonomous Control Engineering
Electrical Engineering Department
University of New Mexico
Albuquerque NM, 87131

Abstract

An algorithm to detect humans in a digital image as the Subject of Interest (SOI) is introduced. This algorithm is based on skin tone detection using fuzzy sets. Color and tint values of each pixel are extracted and used as input to the fuzzy rule base system. The output of the rule base is a skin tone mask that represents the location of SOI in the image. The above mentioned mask is applied to the original image to extract the human SOI.

1. Introduction

One of the most interesting topics in computer vision is image recognition. Several algorithms have been developed to recognize an image using edge detection [6,9], clustering [3,5,7], pattern classification [3], etc. The purpose of these algorithms is to enable a computer to understand its environment from visual information 10]. Among all possible applications of computer vision, the problem of detecting a human SOI is most interesting.

One possible application of *human detection* is to enhance the quality of printed color images which contain human subjects. Once the human subject of interest is detected, the printed image quality can be enhanced based on the characteristics of the color in human subjects [1].

In the following sections a brief review of color images and different image formats is presented. A definition for *skin tone color* is introduced. In the preceding sections a fuzzy decision system is used to compute the degree of skin tone content in each pixel. Finally a mask is generated based on these values using a threshold value and some experimental results are presented.

2. Color Image Formats

The use of color in image processing is motivated because of the human eyes' sensitivity to color. Color is a powerful descriptor that often simplifies object identification and extraction. In order to facilitate the specification of colors some standard color models are used. The most commonly used color model is *RGB* (red, green, blue) . The *RGB* model is widely used for

---

[1] Hughes Space& Communication Company
[2] Currently employed at Intel Corporation

color monitors and a broad class of color video cameras. Other color models are *CMY* (cyan, magenta, yellow) , *YIQ* (the standard for color TV broadcast), *HSI* (hue, saturation, intensity) and *HSV* (hue, saturation, value) . The color model used in this algorithm is *YIQ*, since it separates the color information from intensity. The following section describes *YIQ* model in more detail.

2.1. YIQ Color Format

The *YIQ* model is used in commercial TV broadcasting [4,8]. In this model, Y corresponds to luminance, and *I* and Q are two chromatic components called *inphase*, and *quadrature*, respectively. The principle advantage of using the *YIQ* model in image processing is that the luminance (Y) and color information (*I* and Q) are decoupled.

An *RGB* image can be converted to a *YIQ* format using the conversion below:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

It is assumed that the images are in 24-bit color format and the standard size is 512x480 pixels. In other words, there are 256 (8-bit) distinct colors of red, green and blue associated with each pixel value. This corresponds to a color space of sixteen million colors. Based on this assumption the ranges of Y, *I*, and *Q* are shown bellow:

$$0 \le Y \le 255$$
$$-151.98 \le I \le 151.98$$
$$-133.365 < Q < 133.365$$

2.2. Skin tone definition

After converting the data to *YIQ* format, *skin-tone* pixels can be determined using the following definition:

Definition[2]: *any* pixel in the *Q-I* plane that falls within ± 30 degrees of the *I*-axis and has a normalized magnitude between 25°/0 and 75°/0 is considered a skin tone pixel.

The above definition is equivalent to having a normalized color intensity value between 0.25 and 0.75 and a tint value between 60° and 120°, where color intensity and tint are defined using the following equations:

$$Intensity = \sqrt{I^2 + Q^2}$$
$$\text{Tint} = \tan^{-1}\left(\frac{I}{Q}\right)$$

The resulting skin tone region in the *Q-I* plane is depicted in Figure 1.

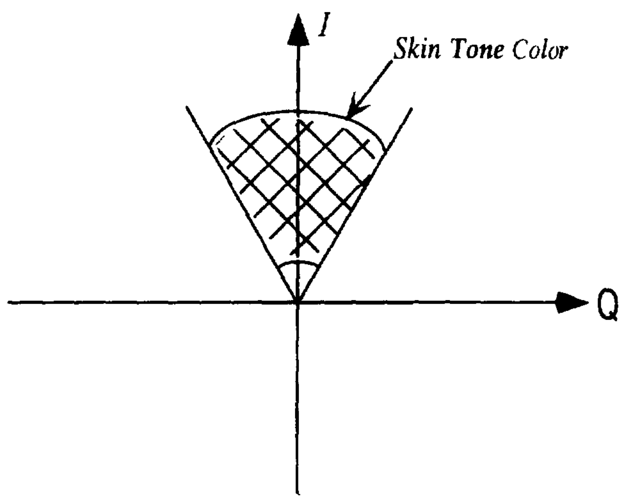


Figure 1: *Skin tone region in the Q-I plane*

3.Skin Tone Detection Procedure

In the above definition the boundaries of the skin tone region are crisp. However, the change of color and tint values in a digital image are gradual. To avoid this inconsistency, color and tint values are fuzzified using the trapezoidal membership functions in figure 3. The fuzzified inputs are then applied to the following rule:

IF *Color is* Skin_Color and *Tint is* Skin_Tint **THEN** *Pixel is* SKIN_TONE

Note that the consequent of the rule is a singleton. Using the above inference, a value between O and 1 corresponding to the degree of skin tone content is obtained for each pixel. These values are then compared with a threshold (λ). Pixels with values greater than λ are assigned the value 1. The rest of the pixels are assigned the value O. The resulting matrix is called a *high resolution mask.* Finally, the subject of interest is detected by multiplying the original image and the high resolution mask. A block diagram of the procedure is depicted in figure 2.

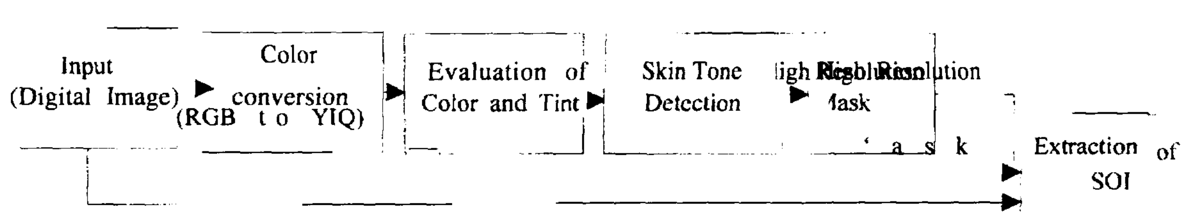


Figure 2: *System block diagram*

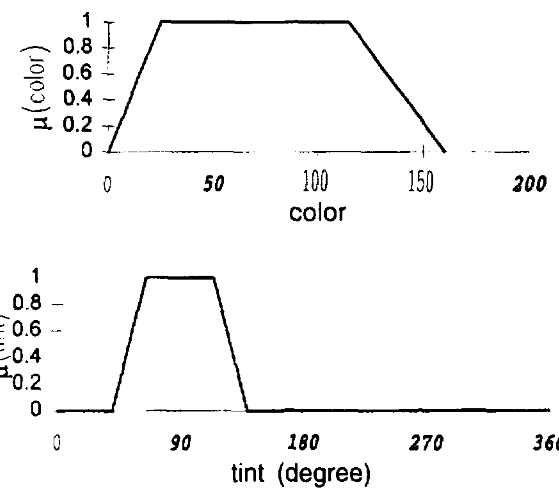


Figure *3: Membership functions for color and tint*

4. Experimental Results

The above procedure was applied to a set of digitized images. The results are depicted in figures 4(a-f). All the images are assumed to have a 512x480 resolution and each pixel has a 24-bit color value. The *RGB* values range between O and 255. Color images are published as gray scale images in this paper due to printing restrictions.

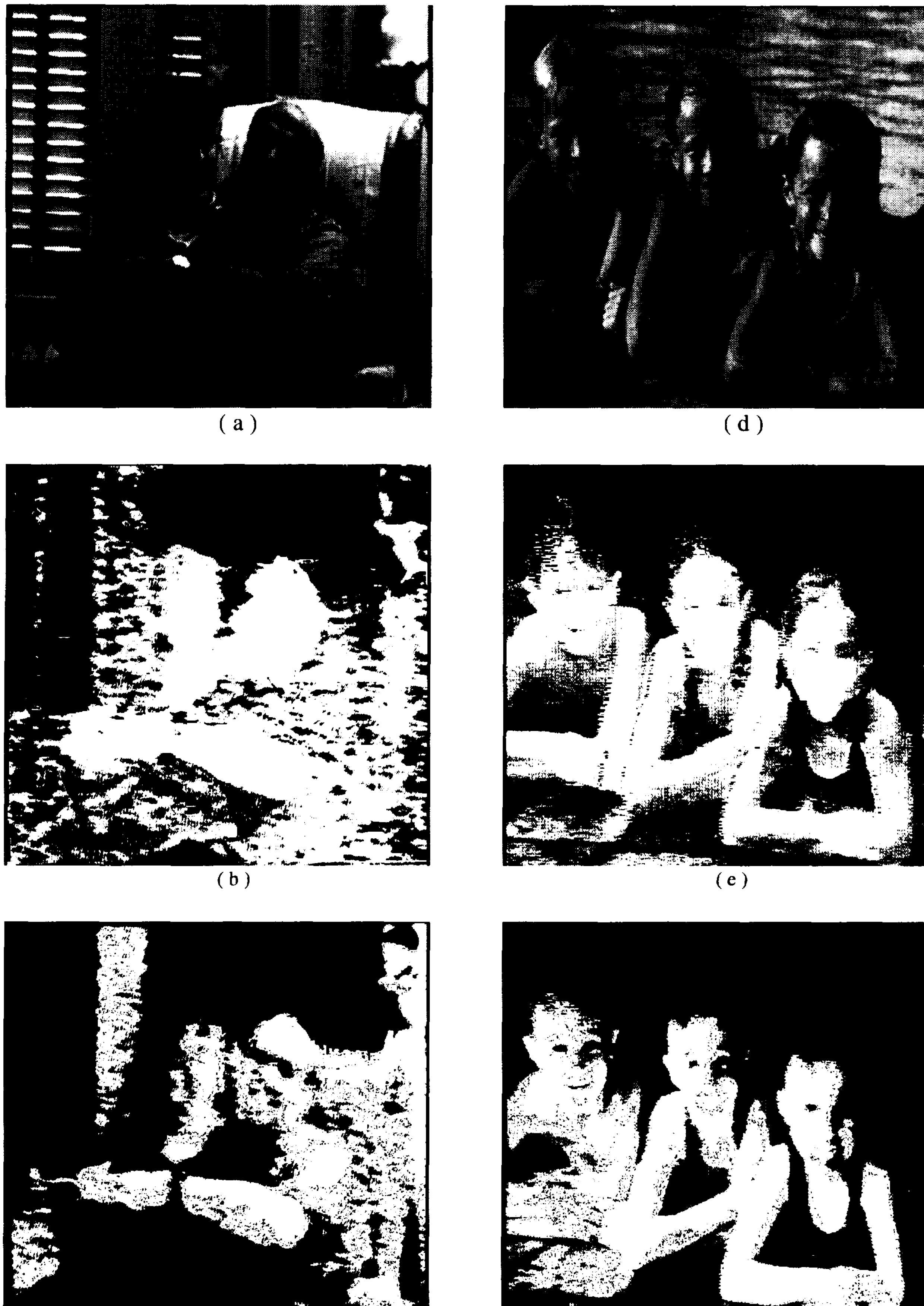
Figure 4(a) represents the original digital image. The result of crisp skin tone detection is depicted in figure 4(b). In this mask, the white pixels represent skin tone. There are a number of non skin tone pixels that are detected as skin tones. On the other hand, the fuzzy logic based mask in figure 4(f) is superior compared to the crisp algorithm. The main concentration of white pixels are now around the faces and the arms of the human subjects. It can be seen that the number of falsely detected skin tone pixels is much less in the fuzzy algorithm than the crisp one. The second test image is shown in figure 4(d). The crisp algorithm's result are shown in figure 4(e). It can be easily noticed that the fuzzy mask in figure 4(f) has less detection error than the crisp mask in Fig(4e).

5. Conclusion

A skin Tone detection algorithm was proposed using a fuzzy decision mechanism. The results were shown to be superior. Further research in this area can be performed to detect other subjects of interest using fuzzy color attributes.

6. Acknowledgments

The authors would like to thank Dr. Mo Jamshidi of the Department of electrical and computer engineering of the University of New Mexico for his kind cooperation and guidance.



(a) (d)

(b) (e)

(c) Figure 4: Skin tone detection Experiment (f)

References

[1] Asgharzadeh, A. and Gaewsky, J., "Applications of Fuzzy Logic in a Video Printer," *Proc. of the first World Automation Congress* (WAC '94), August 1994, published as: *Intelligent Automation and Soft Computing,* Vol. 1, pp 445-448.

[2] Asgharzadeh, A., "Development of a Fuzzy Logic Rule Based Expert System for Polaroid's Freeze Frame Video Recorder (FFVR)," University of New Mexico CAD Laboratory for Intelligent and Robotic Systems, Internal Memo , Nov., 1994.

[3] Bezdek, J. , *Pattern Recognition with Fuzzy Objective Function Algorithms,* Plenum Press New York, 1981.

[4] Gonzalez R., and Woods R., *Digital Image Processing,* Addison Wesley ,1992, pp.225-230.

[5] Lim Y., and Sang U., "On the Color Segmentation Algorithm Based on the Thresholding and the Fuzzy C-Means Techniques," *Pattern Recognition,* Vol. 23 No. 9, 1990.

[6] Pal S. "Fuzzy Tools for the Management of Uncertainty in Pattern Recognition, Image Analysis, Vision and Expert systems," *Int. Journal of Systems Science,* Vol. 22, No. 3, 1991, pp. 511-549.

[7] Ruspini E. , "A New approach to Clustering," Inf. Control, Vol. 15, 1969, pp. 22-32.

[8] Sid Ahmed M., *Image Processing with Applications in Improved Televisions, McGraw-Hill Inc., 1994.*

[9] Tyan C. and Wang P., "Image Processing Enhancement Filtering and edge Detection Using the Fuzzy Logic Approach," *Proc. of Second IEEE Int. Conf. On Fuzzy Systems,* Vol. 1, 1993.

[10] Zhang W, and Sugeno M., "A Fuzzy Approach to Scene Understanding," *Proc. of Second IEEE Int. Conf. On Fuzzy Systems,* Vol. 1, 1993, pp. 564-569.

URC97013

Object Oriented Design and Development of a Computational Mechanics Toolkit

5/3/61

Lance Atencio
University of New Mexico, Department of Computer Science

Walter Gerstle, Malcolm Panthaki, Raikanta Sahu
University of New Mexico, Department of Civil Engineering

Abstract

Object oriented design methods and C++ are becoming popular for software applications of all types. It seems that their popularity has taken longer to reach the realm of scientific and engineering applications. We are currently using these tools in the design and development of a Computational Mechanics Toolkit (CoMeT) to try to benefit from their claimed reduction in maintenance and increase in reusability and extensibility. We hope to achieve these benefits while producing an application that will fill the void between CAD and Analysis.

Benefiting from Object Oriented Design

in the past, most scientific and engineering software applications were designed using a procedural approach and implemented with a compiled programming language like FORTRAN or C. The design and coding for these projects were usually done by the engineers who had defined the problem specification. Unfortunately, these engineers rarely had any formal training in software design or programming standards, which led to code that was very costly to maintain and difficult to modify. This problem is significant because maintenance has been shown to be the largest cost during an application's lifecycle (Johnson and Foote 1988).

Technical applications are beginning to be developed using object-oriented design and the C++ programming language for several reasons including the reduction of high maintenance costs. In an ideal development environment of this type, engineers and scientists from the appropriate field determine the specifications for the application, software engineers design the product, and programmers implement the design. When the ideal situation can't be met, it is important that the developers, whatever their background, be trained in software design and use coding standards to achieve a product that will be maintainable.

The primary benefits of object-oriented design are claimed to be reusability, understandability and maintainability (Johnson and Foote 1988; Headington and Riley 1994; Somerville 1996). Better reuse is achieved through modularity, information hiding and inheritance. Well designed objects are more easily reused through increased modularity. Information hiding provides a barrier between an object's interface and its implementation, giving improved maintainability; a change in an object's implementation will not lead to changes in other classes because its interface can stay the same. Inheritance provides the ability of new objects to be extensions or slight modifications of existing ones. The developers gain a deeper understanding of the levels of abstraction in a design by focusing on

the design of the objects.

In practice it is difficult to achieve code that is easy to maintain even when using an object-oriented approach (Johnson and Foote 1988). Often people confuse the use of C++ and other tools as the answer to their maintenance problems. C++ is a programming language and does not produce object-oriented code on its own. Object-orientation is achieved by design. The key to better code is careful design. Only when good design is combined with a clearly defined development process and good coding standards can a software product fully benefit from the object-oriented approach.

Design and Description of CoMeT(Computational Mechanics Toolkit)

CoMeT is a software framework being designed using an object-oriented approach (Wirfs-Brock et al, 1990). It is intended to be used as a modeling and analysis environment for numerical simulation of multi-physics natural processes. It has a structure that allows the modeling and analysis of problems in which both the mathematical model and its numerical discretization can change during the analysis. Some general-purpose commercial finite element programs currently allow parametric descriptions of geometry but, to our knowledge, none allow significant modeling aspects, such as geometric topology or material model, to change as an analysis progresses. Unlike other finite element programs, which are designed to analyze only a pre-specified finite element mesh, CoMeT has a structure in which a model can change state in significant ways not previously allowed. For example, a discrete crack may develop and propagate automatically as an analysis proceeds, or the number of holes in a part may change during a design optimization.

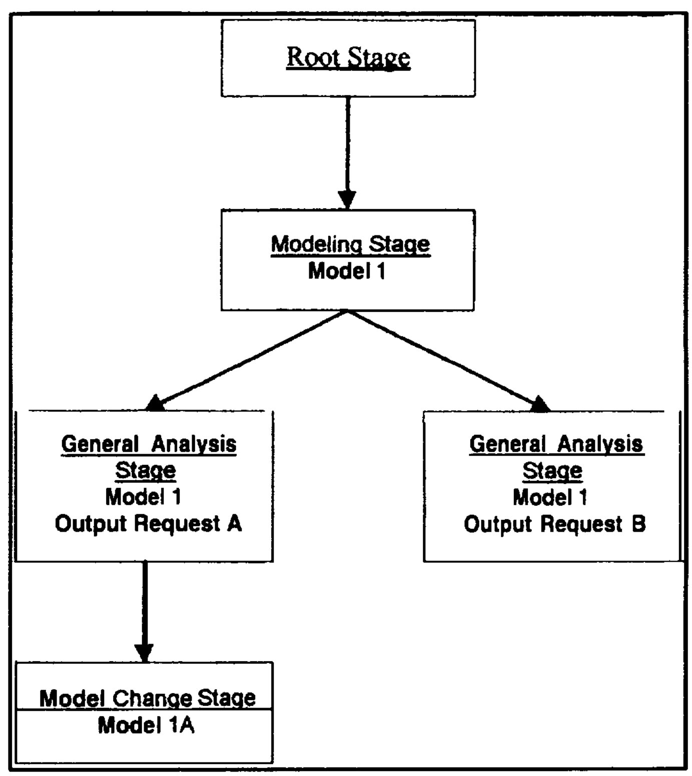


Figure 1
Project Tree

CoMeT is designed with a virtual geometry interface (VGI) which can be adapted to various geometric modelers. Currently, the geometry is represented by the ACIS geometric modeling software package (Spatial Technology, Inc. 1996). Modeling and analysis performed by users are maintained as projects in a tree structure. This structure is designed as a composite of "stage" classes of various types for creating, editing and analyzing a "model" class, as shown in Figure 1. The project tree is viewable graphically to provide an interface for management of the project, The project tree provides the means to create and maintain variations of a model in different stages. The project tree structure allows parallel (by different

numerical analysis applications) or sequential analyses to be performed on a model. Parallel stages can be compared to each other in parametric studies, while sequential stages allow evolutionary analyses, such as crack propagation, to be conducted.

CoMeT will typically operate as follows. A project tree is initiated in which a mathematical model of the problem is created. Feature-based parametric geometry is generated for the model. The geometry can be imported from a CAD application or created using CoMeT. Problem boundary conditions are added to the model as required. A numerical model may then be automatically or semi-automatically created. The numerical model can then be routed to a numerical analysis application along with any required steering conditions. Any number of studies of the mathematical and numerical model may be performed. Results data returned from the numerical analyses are viewable graphically along with an updated (if necessary) mesh and geometry. The resulting feature-based parametric geometry can be exported to a CAD application (contingent upon a standard protocol for exchanging feature-based parametric geometry). Thus, it is hoped that the design-analysis loop will fin-ally be closed as illustrated in Figure 2.

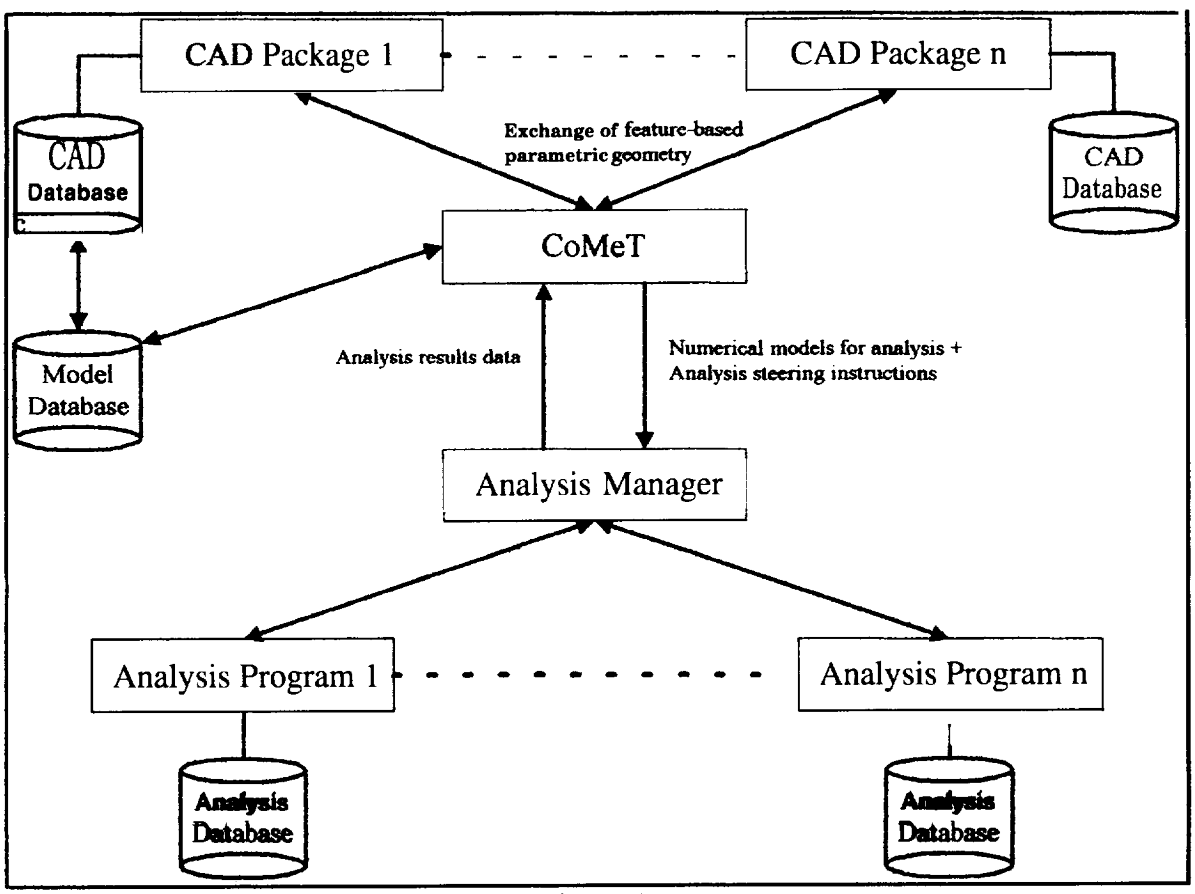


Figure 2
CoMeT's place in design and analysis

Architecture of CoMeT

The atomic unit of the first level of abstraction in the architecture of CoMeT is what we call a Primary Component (PC). This is the highest level view of a complex system and provides ease of understanding as well as the first level of modularity. A PC also serves as the unit of distribution of tasks across a network of computers (e.g. in a UNIX environment the PCS of CoMeT could be separate processes). The PCs of CoMeT include Command Source, Session, Job and Database Server as shown in Figure 3. A connection between two PCs indicates that communication can occur between those PCs. Data redundancy is minimized across PCS by design. Each of the PCs are divided into subsystems. Within the subsystems are the classes that define CoMeT.

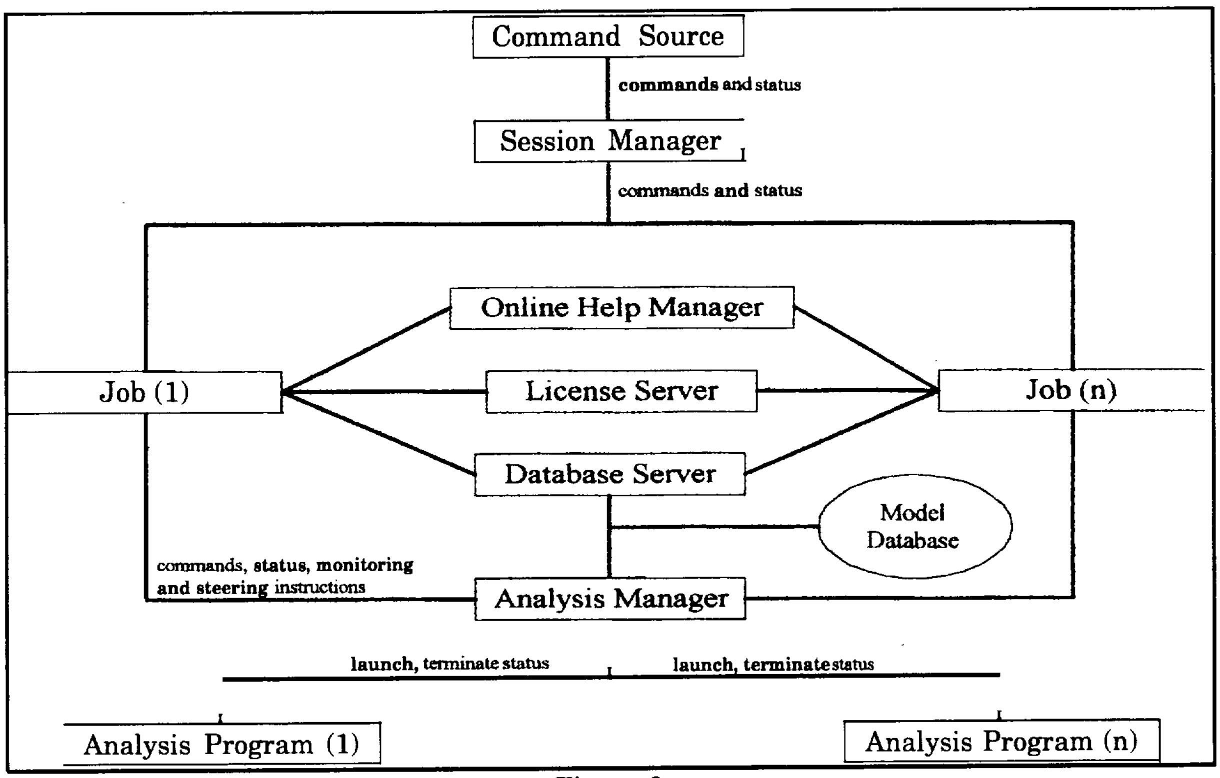


Figure 3
CoMeT Primary Components

The Command Source PC generates commands either through the Graphical User Interface (GUI) or a script file. This allows execution of CoMeT to be controlled either interactively by user input through the GUI or in batch mode using a script file. Regardless of the source of a command, the Command Source routes it to the Session PC.

A session of CoMeT can consist of one or more user created Jobs running concurrently. At any given time there is only one current Job. Commands from the Command Source are routed by the Session PC to the current Job, Each command is executed by a Job. As each Job is a sperate process, the existence of multiple Jobs in a session allows for the

parallel execution of multiple, sequential streams of commands (one per Job).

The Database Sewer PC provides access to and modification of model data. Multiple Jobs can access and/or modify this data simultaneously. The Database Server will provide the necessary transaction management to prevent conflicts and the corruption of data.

CoMeT Data Objects

Within the Job primary component are subsystems that contain the objects which are used to describe the model and analysis of a numerical simulation problem. These objects are called the CoMeT Data Objects (CDOs) because they maintain all the data required for a problem being modeled in CoMeT. Static relationships between CDO classes are maintained in an entity relationship class which is used to control how aggregate hierarchies are built. The run time relationships between instances of CDOs are maintained in a directed acyclic graph (DAG) which facilitates generalized queries for the current relationships any given CDO.

The CoMeT data objects are the focal point of the application and will maintain all the data that makes up the database for the application. In order to achieve the benefits of object oriented design, it is essential that the CoMeT data objects be designed correctly and provide the means to capture the details of all possible model requirements.

Example Problem

A better understanding of how CDOs relate to a computational mechanics problem can be obtained through an example. Consider the plate with a hole in the example problem shown in Figure 4. It has three prescribed conditions which are modeled by PrescribedCondition CDOs. Each of these is related to model geometry (vertices in this case) by PrescribedConditionMediator CDOs and a GeometricDomain CDO. The plate is modeled by an Assembly CDO consisting of one Part CDO.

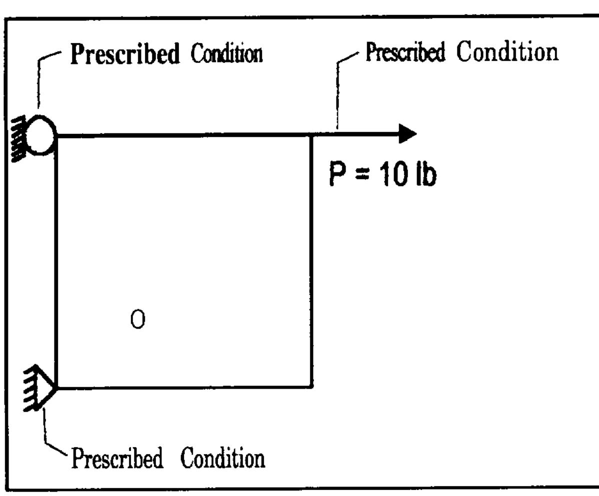


Figure 4
Example Problem

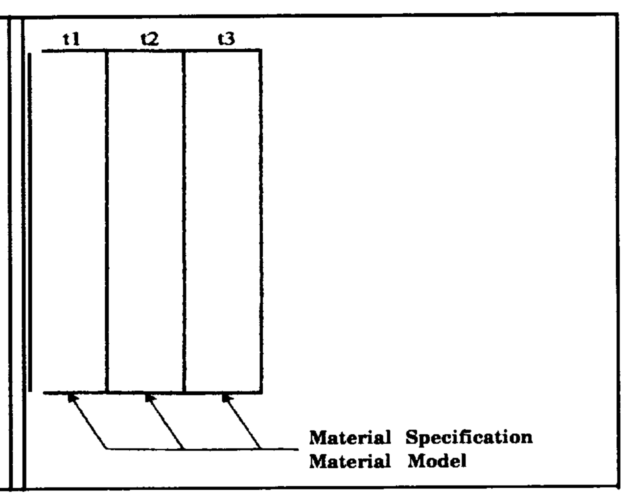


Figure 5
Section

Similarly, the example problem must have RefGeometry, GeometryEntity, and Mesh CDOs to describe its physical shape and numerical model. It also has a Section CDO to model its cross section (made up of three different layers) as shown in Figure 5. MaterialProperty CDOs will be associated with the layers of the Section CDO either through a SectionGeometryMediator CDO with a MaterialSpecification CDO and a MaterialModel CDO, or through a PreIntegratedProperties CDO.

Additional Features

CoMeT provides additional features for the modeling and analysis of a problem. There will be data libraries which can be shared within and between project trees. A library contains objects that may be repeatedly used, such as geometry, boundary conditions, material specifications, analysis controls, etc. Any type of CoMeT data object (CDO) may be saved in a library. The ability to monitor and steer analyses in progress is another feature that is planned for CoMeT. This will allow a user to stop an analysis that has gone awry and restart (possibly at a previous stage) with an adjusted numerical model. There is also an abstract type of object that encapsulates all "quantities" that are used in the program. Through extension of this object we plan on having fuzzy numbers which will be the basis for the ability to perform fuzzy finite element analysis.

By combining the compiled language, C++, with the interpreted and extensible functional language, Scheme, we are developing a command language for computational mechanics which will be at the core of CoMeT. Low-level atomic Scheme commands are implemented in C++, while higher-level commands are implemented in a Scheme script and interpreted at run time. In this way, diverse computational mechanics software applications can be easily written in Scheme using a relatively small set of Scheme atomic commands which are implemented in C++. As new atomic commands are required, they can be implemented, as Scheme is an extensible language.

The environment in which CoMeT is currently being developed has been custom made by the designers of CoMeT. It provides the ability for multiple users to create, edit and test files in their own directory structure while maintaining a central repository of all the working revisions of completed code. The versioning is based on the Concurrent Versioning System software package (Free Software Foundation, Inc., 1996). Additionally, we have created a build system that is made up of script and executable files that are tailor made to the directory structure created for CoMeT.

Conclusions

Our development team is in an academic environment and consists primarily of civil engineering and computer science graduate students. We have been in the design process for the last two years with about six months of overlapping implementation time. We have found implementation in C++ to be a lengthy process (most likely due to the inexperience of our team in the implementation of large programs in C++) and expect our first prototype to be completed in the Spring of 1997. We will not be able to see how well the object oriented design of CoMeT achieves the benefits of reuse, extensibility, and maintainability until it is completed and starts being used. We have, however, already seen how using an object

oriented approach has lead to much greater thought about, and subsequent understanding of, the abstract ideas and theory required for finite element analysis than has been experienced by developers using procedural types of design.

We have seen success in our custom made development environment. It has proven _ itself by repeatedly meeting the required objectives. Over half a dozen developers have been using the environment with satisfactory performance for several months.

References

Free Software Foundation, Inc., (1996) *CVS.* Boston MA. http://www.fsf.org/fsf/fsf. html

Headington, M. and Riley, D., (1994) *Data Abstraction and Structures Using C++.* Lexington MA: D. C. Heath and Company.

Johnson, Ralph E. and Foote, Brian, (1988) *Designing Reusable Classes.* Department of Computer Science, University of Illinois.

Somerville, I., (1996) *Software Engineering, fifth* edition. Reading MA: Addison-Wesley.

Spatial Technology, Inc., (1996) *ACIS.* Boulder CO,

Wirfs-Brock, R., Wilkerson, B., Weiner, L., (1990) *Design of Object-Oriented Software.* Englewood Cliffs, NJ: Prentice-Hall.

Page intentionally left blank

URC97014

The V experimental visual programming language

Mikhail Auguston
Alfredo Delgado
Computer Science Department, New Mexico State University
Las Cruces, NM 88003
email: {mikau, adelgado}@cs.nmsu.edu

S,4/61

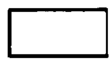
The V language design is an experiment with a visual representation of dependencies between data and processes. Data-flow diagrams are most commonly used to represent those dependencies in visual programming languages, for instance, in LabVIEW [Baroth, Hartsough 951. and Prograph [Cox, Gilles, Pietrzykowski 951. In this paper we consider such data structures as sequences (vectors), and 2-dimensional matrices. The aim of our work is to demonstrate how iterative control constructs typically used for sequence and matrix traversal could be visualized and adapted into data-flow paradigm. We expect that this language will be used as a rapid prototyping tool for experiments with algorithms. The following ideas have contributed to our design.

- The program is rendered as a two-dimensional data-flow diagram that visualizes the dependencies between data and processes. The diagram defines the order of function calls and the data dependencies between function calls.
- Data-flow diagram supports the possibility of parallel execution of threads within the diagram. This approach to Visual Programming Language design became quite common in recent years, see e.g. [Baroth, Hartsough 95], [Glaser, Smedley 95], [Kimura 951.
- The data-flow paradigm is closely related to the functional programming paradigm [Bird, Wadler, 88]. The main difference is that a diagram may have several output ports, while a function returns a single value.
- Diagrams maybe nested, and actually are similar to the notion of procedures in common programming languages. Diagram calls maybe recursive.
- V introduces special loop constructs for sequences. matrices, and multisets. Here we've followed an approach suggested in the SequenceL language [Cooke 961.
- The use of a pattern mechanism for vectors and matrices eliminates the need for explicit index variables and makes the position of adjacent items within an aggregate visible and easy to understand. It is especially useful in the case of two-dimensional objects (matrices). Patterns also introduce temporary names for items.
- Pictograms and text can be combined together. Simple expressions could be better rendered as a plain text.
- In many cases an iterative algorithm description for vectors and matrices is preferable to recursive descriptions. The V language allows to benefit from combining iteration with recursion.
- Ellipses used in the iterative patterns make the iteration description more visible and comprehensive by providing the direction of iteration. The use of ellipses also solves the problem of nested iterations, e.g. within a matrix.
- Two-dimensional notation makes it possible to introduce such useful control structure as iteration synchronization. This provides an easy to understand way to describe operations that involve multiple vectors or matrices.

- The usage of two-dimensional pictograms reduces the number of lexical and syntactical elements needed to write the program. Good picture may be worth tens of lines of linear code.

Data-Flow Diagram Notation

Programs in V are rendered as data-flow diagrams that contain pictograms representing data and processes. The following pictograms have been used in the rest of this paper.



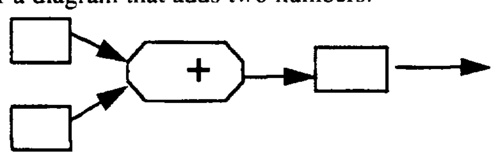
value box, denotes a value (scalar or aggregate)

operation box

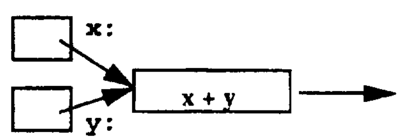
Semantics. An operation box has input and output ports to connect the operation with its input and output data. An operation box fires when and only when the following conditions are satisfied.

All input values are delivered to the corresponding input ports and all output values produced at the previous execution cycle are consumed by input ports of the connected nodes downstream. In the case of user-defined operations (diagrams) there are not more active nodes within the diagram (i.e. all previous computations are over). There never is more than one data item in the channel between the sender's output port and a receiver's input port. It is convenient to draw data flow in the diagram from left to right.

Here is an example of a diagram that adds two numbers:

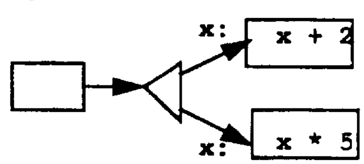


The previous example can be presented in the abbreviated form by providing names at the input ports.



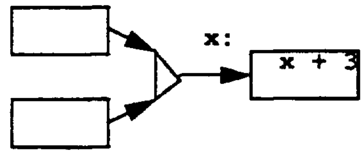
Input values denoted by x and y are added using predefine operation '+'. Box expression represents the result of an expression evaluation. Names defined at the input ports are visible only in the immediately following boxes.

The *fork* pictogram receives a single value and passes several copies of this value downstream.



Semantics. The fork node fires when it's input port receives a value and all output values from the previous cycle are consumed by input ports downstream.

The *merge* node *fires* when at least one input port receives a value and passes it to the output port.

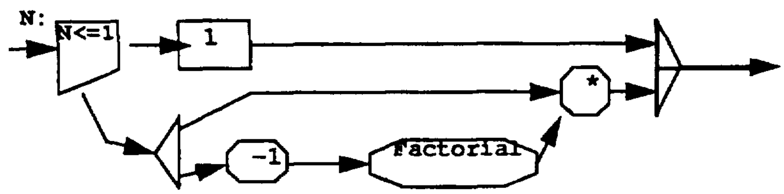


Semantics. The **fairness** property guarantees that each input value will be processed.

In order to be able to draw diagrams with branches and loops we need a conditional flow switch. A conditional switch has several input ports (at the left side of the pictogram) and the same number of output ports 011 each of "True" and "False" sides. The Boolean Expression is evaluated and the flow of input values is switched to the corresponding side.

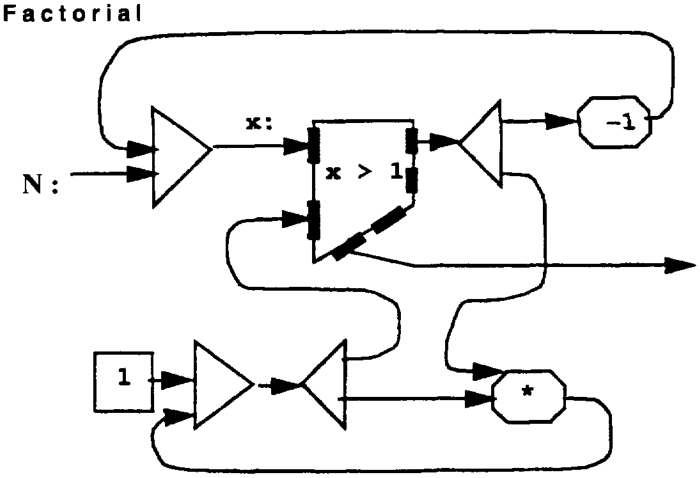
Example. This is the most common recursive version of the Factorial program.

Factorial



Example. A diagram that computes a factorial with an attempt to parallelize some threads.

Factorial

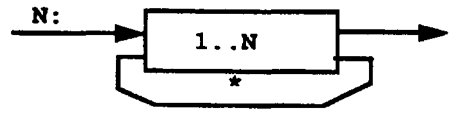


Regular Computations

The V language supports regular computations, such as applying operations +. *, MIN. MAX, && (Boolean AND), || (Boolean OR) to the whole vector or matrix, or along some dimension within a matrix.

Example. Yet another way to describe factorial function in our formalism is as follows:

Factorial



The 1..N expression yields a sequence of integers from 1 to N.

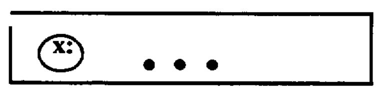
Iterative Patterns and Synchronization

An **iterative** pattern provides temporary names for the values associated with the current item and the index value of the current item. The iterative pattern also defines the order of iteration performed on a vector or matrix.

a pattern that matches a single value (scalar or aggregate)

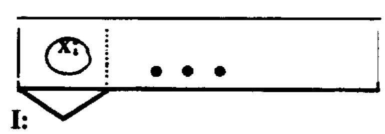
a pattern that matches a group of values (a sequence within a vector, a row or a column within a matrix, or a submatrix within a matrix)

A name followed by ':' and placed inside a pattern box denotes the value or the set of values within this box. The iterative pattern describes the direction of iteration.

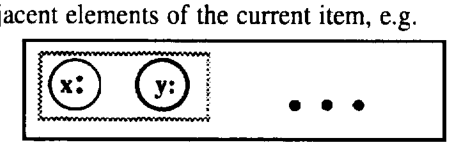


This pattern defines an iteration over the vector from left to right (from the items with the smallest index values towards the largest index values), The current iterative item is given a temporary name x, which is visible downstream of the diagram in operation and data nodes immediately connected to the pattern node.

It is possible to provide a temporary name not only for the iterative item value, but for it's position (index value) as well. In the following example the index value of the current item can be referred to as I. The dotted line is used to indicate to what pattern element the index pattern is related.

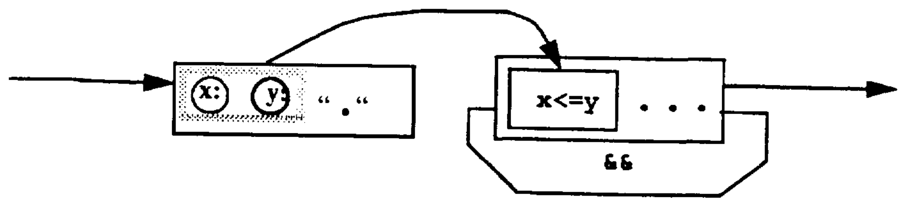


The pattern may involve some adjacent elements of the current item, e.g.

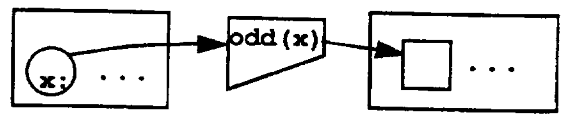


Two data boxes containing iterative patterns and connected in the diagram are *synchronized. This* means that items in each of them are visited in the same order. As a rule the value of the second synchronized aggregate is constructed from the value of the first aggregate.

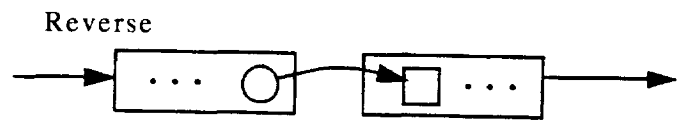
Example. To check whether a vector is sorted in ascending order.



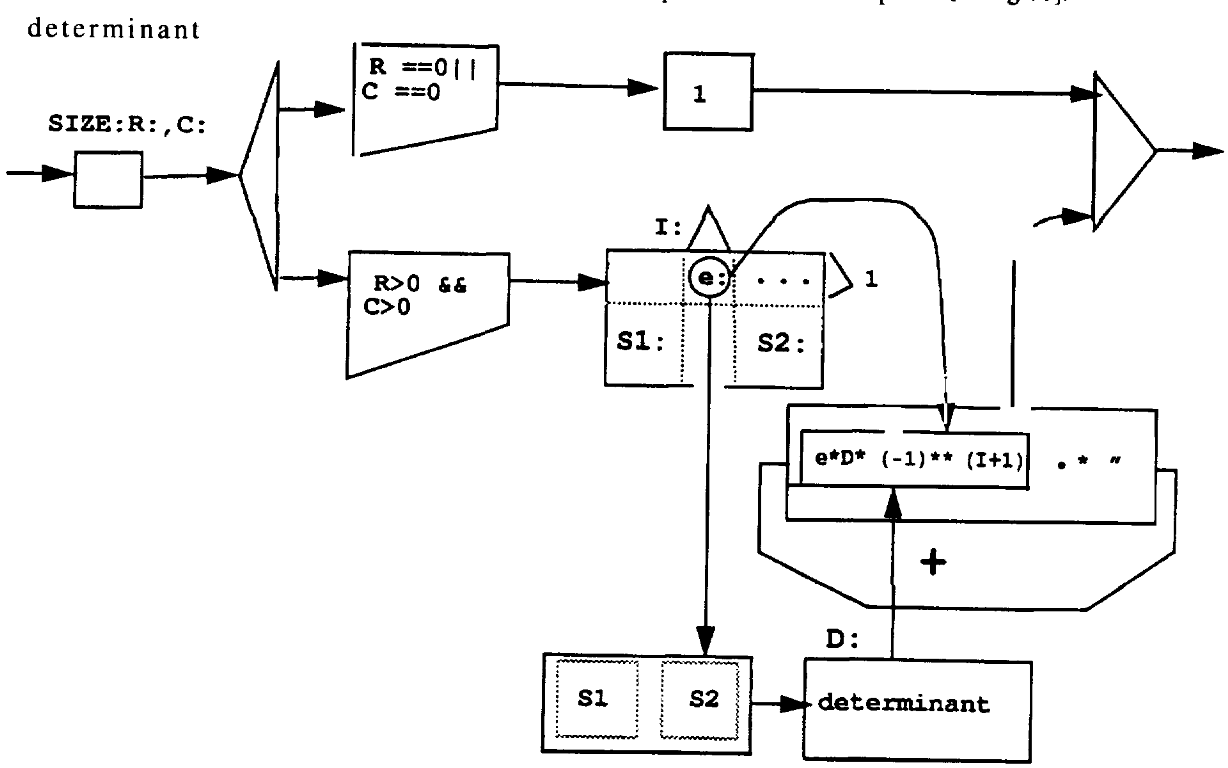
Example. Select all odd numbers from a vector.This diagram is an analog of the list comprehension operator in functional languages such as Miranda [Bird, Wadler, 88].



Example. To reverse a vector.



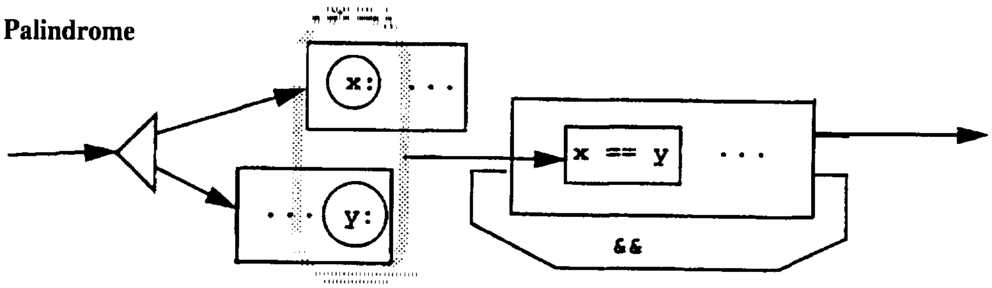
Example. Diagram to calculate matrix determinant. This example follows au example in [Yeung 88].



The pattern may involve more than one aggregate in the diagram. In this case the iteration over multiple aggregates maybe synchronized. The dotted frame denotes the virtual group of objects which are synchronized during the iteration. The number of iterations is determined by the number of items in the shortest aggregate.

Example. checking whether a vector is a palindrome,

Palindrome



Prototype Implementation Effort

The work on the V visual programming language has been supported during Summer 1996 by PACES (Pan American Center for Environmental Studies) and NASA. In particular, the visual programming interface has been designed for the SequenceL language [Cooke. 1996]. The prototype implementation of the V language is underway by Alfredo Delgado and Shridhar Bidigalu. The JAVA programming language is used to design a graphic editor and run-time environment for V. This will ensure portability of the V environment to UNIX, Windows, and MAC OS platforms.

References

[Baroth, Hartsough 95] E. Baroth, C. Hartsough, Visual Programming in the Real World, in Visual Object-Oriented Programming, Concepts and Environments (cd. M. Burnett, A. Goldberg, T. Lewis), Maiming 1995, pp.21-42.

[Bird, Wadler, 88] Bird R., Wadler P.. Introduction to Functional programming, Prentice Hall, NY, 1988.

[Cooke, 96] Cooke D. An introduction to SequenceL: a language to experiment with constructs for processing nonscalars, to appear in the Software Practice&Experience, 1996

[Cox, Gilles, Pietrzykowski 951 P.T. Cox, F.R. Gilles, T. Pietrzykowski, Prograph, in Visual Object-Oriented Programming g, Concepts and Environments (cd. M. Burnett, A. Goldberg, T. Lewis), Manning 1995, pp.45-66.

[Glaser, Smedley 951 Glaser H., Smedley T., PSH - the next generation of command line interface. in Proceedings of the 11th International Symposium on Visual Languages, VL'95, IEEE Computer Society press, 1995, pp. 29-36.

[Kimura 95] Takayuki Dan Kimura, Object-Oriented Dataflow, in Proceedings of the 11th International Symposium on Visual Languages, VL'95, IEEE Computer Society Press, 1995, pp.180-186.

[Yeung 88] Ricky Yeung, MPL - A Graphical programming Environment for Matrix Processing Based on Logic and Constraints, in IEEE Proceedings of the Workshop on Visual Languages, 1988, pp. 137-143.

URC97015 5/5/31

A Biofunctional Approach to Decision Making

Jalal Baghdadchi, Graduate Research assistant,
North Carolina A&T University,
Department of Electrical Engineering

Abdollah Homaifar, Associate Professor,
North Carolina A&T University,
Department of Electrical Engineering

Ali Iran-Nejad, Professor,University of Alabama,
Department of Educational Psychology

Keywords: whole-theme, biofunctionality, associationism, connectionism, cognition, OBA, MCF, active regulation, dynamic regulation, inhibition

Abstract

Decision making is the core of any control system. The structure of traditional decision makers has been based on a knowledge-base composed of a set of IF-THEN rules, and a sensing mechanism. The driving metaphor for these types of decision makers is the existence of an explicit memory base and its ability to analyze and express by a set of mutually disjoint rules, any situation that the controlled plant experiences. Following the fact that certain situations in real-time do not lend themselves easily to this kind of partitioning, we are proposing a decision making unit which confronts the problems in their entirety, thus, functioning to a great extent like the human brain.

1. MOTIVATION

In the engineering world, when encountering a problem, we normally tend to try to partition the problem into smaller pieces, deal with individual pieces then put the whole thing back together. This strategy is referred to as the analytical method, however this not necessarily the best way of dealing with a problem. Not all the real life issues lend themselves easily to this type of segmentation, Often, in trying to divide a problem into pieces we forget the glue which holds these pieces together. Suppose we want to teach someone how to ride a bike. We can give the individual some insight about the pieces of a stationary bike such as the handle bar, pedals, braking mechanism, etc.. Of course chances of the individual becoming able to ride a bike are not very good. Obviously, unless there is motion the knowledge about the pieces of the bike are not going to be helpful in getting the individual to acquire the ability to ride a bike. On the other hand, if we give the potential bike rider a bike and have him/her just get on it and try to go forward by pedaling, chances of success will be considerably higher (Iran-Nejad et al., 1995), The first approach is the analytical method, the second is a whole theme approach. In whole theme approach we try to approach the problem as a single item as opposed to analytical method, where we try to divide the problem to pieces

2. OBJECTIVE

The objective is to come up with an improved decision making unit which is modeled after the functioning of the human brain and is applicable to real-time engineering problems. We will first survey the existing cognitive theories of knowledge, then put together our biofunctional model.

3. COGNITIVE THEORIES OF KNOWLEDGE:

For the real time decision maker we will be dealing with sensing, storing and retrieving of information. The equivalent terms for these activities in the field of psychology (cognitive theories of knowledge) are cognition, learning and remembering. Notice that the equivalency is approximate. For example, in the case of sensing versus cognition, cognition goes beyond simply sensing of the information. In fact, cognition is sensing, processing, and possibly coming to a conclusion. There are three major families of theories attempting to describe the functioning of the human brain. These are Associationist theories, Connectionist theories and the Biofunctional model, The first two rely on explicit memory as the knowledge base. Biofunctionality is a whole theme approach. We will

consider Information Processing Theory as an example of associationist theories, and Adaptive Resonance Theory along with neural Networks as examples of connectionist theories. Biofunctionality will be discussed in detail.

3.1. Information Processing Theory (IPT):

Information processing theory is historically the first major theory to attempt to model the functioning of the human brain and the learning processes. It is based on computer metaphor. IPT considers the human brain a storage/retrieval structure. The backbone of the system is three separate types of memory: sensory register (SR), short term memory (STM), and long term memory (LTM). The set-up of the system is very similar to a microprocessor hardware line-up (with registers, RAM and ROM), The capacity limitation associated with explicit memory, suggests that the learning should take place sequentially (Iran-Nejad et al., 1995). Prerequisite parts of knowledge have to be internalized first, before they can be combined to construct more complex knowledge.

Learning Mechanism: Learning takes place in three hierarchical phases: cognitive phase, associative phase and autonomous phase. During cognitive phase, the learner mindfully gathers isolated pieces of information. A declarative storehouse of concepts, facts, and principles is built. During the associative phase, connections between the stored pieces of information are made and strengthened, The declarative representation gradually changes to procedural representation. Finally, in the autonomous phase, procedural skills become fully automatic and acquire the ability to self-trigger. In other words, during this phase, no executive control or mindful allocation of attentional resources are required.

Limitations of IPT: The backbone of IPT model is a system of explicit memory structures (SR, STM, LTM), which implies capacity limitations. Capacity limitations in turn, imply that the learning process is sequential. Real learning is both sequential and parallel. In fact one major advantage of the human brain over man-made brain-like systems, is its ability to deal with simultaneous incoming of enormous amount of inter-related data. Furthermore, the IPT model, casts the brain as a predominantly passive memory structure, run by a mindful active control. Obviously, the human brain packs much more capability than a static storehouse. Finally, the IPT suggests that learning is intentional. The real learning is of course both intentional and incidental.

3.2. Adaptive Resonance Theory (ART):

The human brain receives and processes information both serially and in parallel. Consider a picture of someone who you know. If the picture is shown to you, you would not have much of a problem realizing whose picture it is. On the other hand, if the same picture is presented piece by piece, one piece at a time, one would have an awful hard time recognizing the fellow in the picture, This demonstrates the fact that certain categories of knowledge are best perceived through parallel processing. Learning is based on perception matching; "Our perception is often matched against our expectations" (Grossberg, 1995). Expectations lead to focus of attention to data worthy of learning. The need for parallel processing capability and context dependency of learning, as well as the brain's capability to work backward in time (hence the idea of resonance), calls for a more flexible learning model. One such model is Stephen Grossberg's (Grossberg, 1995) adaptive resonance theory.

ART model set-up and learning mechanism: ART model structure is based on two memory elements and an inhibition mechanism. The two memory units are long term (LTM) and short term (STM) memories, Short term memory has two components: feature STM and category STM. Learning is based on perception matching. The input, or the detected feature is housed in the feature STM. LTM dispatches a match for the sensed feature and directs it to be placed in the category STM. This match, is the recognition category or learned expectation. Next, the sensed feature and the learned expectation (match, or recognition category) are compared. If the match is reasonably close, a conclusion is made. If there is a mismatch, the cycle is repeated. Inhibition is used for selective amplification and suppression. If a sensed pattern closely resembles an expectation, the inhibiting systems are turned on (focusing attention). If there is a mismatch the inhibiting mechanisms will be weakened or turned off, thus, allowing further search (Grossberg et al., 1994). "Whether or not resonance develops, depends on the level of mismatch, or novelty, that the system tolerates."

Limitations of ART: From practical point of view, ART is preferable to IPT in the sense that it compensates for some of the short comings of IPT. The idea of inhibition however, is a weak point for ART learning model. Inhibition as explained in (Grossberg, 1995), is triggered when the system realizes that its previously initiated perception matching is leading to a mismatch. This is analogous to accelerating a vehicle from a constant speed,

then applying the brakes to slow it down to the original speed. No useful result is obtained as a result of this action/reaction pair, but both use the system energy and drain its processing capability.

3.3. **Neural Networks:**

The neural networks are modeled after the human neural system. The basic elements of the model are neurons and nerves. Nerves provide the connections between the central control and the organs. There are two neural subspaces filled with neurons. One subspace presents the input or sensing environment. input pattern(s) is (are) received via these neurons. Output appears on the output subspace. Neurons of the input subspace are connected to neurons of the output subspace via nerves (connections). Knowledge (rules) is housed in the weighted connections between input and output neurons.

Learning Mechanism: As mentioned above, the knowledge of a neural network is stored in the network of its connections. "Learning" amounts to assigning of weights to these connections. Initially an arbitrary (random) set of weights are assigned to the network. Then, through trial and error the weight assignment is refined. The connection weight set which maps the input (such as an object to be recognized) as closely to the ideal output (an expected image) as possible, is the final form of the knowledge retained.

Limitations of Neural Networks: Just like the associative theories discussed in the previous sections, the neural networks are susceptible to the capacity limitations of explicit memory structures. Also, the success of the system in a cognitive mission depends on the design (mindful monitoring of the initial] y stored data). The neural networks are capable of one feature detection at a time.

3.3. **Biofunctional Model:**

Before we explore the biofunctional model, let's consider a couple of biological examples. Assume you wish to pick up an object using your hands. If you are right handed, like I am, your right arm will move towards the object in an attempt to pick it up. The neural network model suggests that there is a nerve connection between the brain (or spine) and your right hand. The command to pick up the object travels through this nerve from brain to the right hand.

Biological experiments however, reveal that the response of a certain organ to a command does not depend on the existence of a physical connection between the organ and the command issuing center (brain or spine). The experiment with the trained salamander clarifies the point. In this experiment, a salamander was trained to push a little door to let loose his food. The salamander normally used his right front leg (hand) to push the door open. When the trainer tied salamander's right front leg to his body, he used the left hand to push the door open (Homaifar et al., 1995). When both front legs were tied, the salamander resorted to using his head in an attempt to push the door open. Had the animal's reaction to command been dependent on the existence of a physical connection from brain to his right hand, he would not have attempted to execute the routine using alternative organs.

Another biological example to the same effect as the point made above, is the blood circulation system. Blood carries different types of particles, hormones, and cells throughout the circulation system. Every organ absorbs the material relevant to itself and lets everything else flow by. Somehow the organs have the distinguishing capability to target and absorb their relevant material only.

The two examples above suggest that the commands, news and information travel throughout the system and that biological activities (such as a certain organ performing a task) do not need a physical connection to/from the brain in order to receive the command. In other words, the awareness is universal. Once an event takes place, the news is available throughout the system. A learning model accommodating this important feature of biological entities is closer to real life and suits our purpose better.

Biofunctional Model: *The* core of the biofunctional model is the existence of two types of brain activities. Learning is achieved as a result of the interaction between these two activities. The two activities are on-going-brain-activity (OBA), and momentary constellation firing (MCF). OBA is an unbounded, ever present, ever evolving and ever rearranging contiguous body of knowledge. MCFs are short lasting appearance (detection of or

reaction to) events.

A learning system includes sources of control, functionally autonomous subsystems (organs), on-going-brain-activity (OBA), and the capability of generating MCFs. The subsystems can be thought of as bodily organs. OBA is the main (and only) body of knowledge. Note that OBA is not static. MCFs on the other hand, are momentary phenomena. At any given instant of time the system may or may not experience the appearance of any MCFs. We will further elaborate on OBA and MCFs through the following analogy: Consider a highway traffic arrow (of the kind which are usually used while construction goes on). The arrow is composed of many light bulbs. If all the light bulbs are on or off, we will see the shape of an arrow, If light bulbs go on and off with a certain pattern, we will experience the illusion that the arrow is pointing to a certain direction. The original arrangements of the light bulbs (independent of their on / off status) is analogous to OBA. The turning on and off of each individual light bulb, can be thought of as MCFs. The combination of the background provided by light bulbs, and their going on and off result in the learning experience (the arrow pointing to a direction).

OBA and MCF in action; Two examples: Consider yourself as the learner and someone who you know and regularly interact with as the subject to learn about. Let's define the summary of all your information about this individual, that is your opinion, as the OBA. Two months from now you may have a different opinion about the individual. The change in the previous opinion will then be caused by the new experiences and interactions that you have had with the individual. These experiences are the MCFs. The relevant experiences (MCFs) will change the composure of the previous OBA (rearrange it).

Again consider yourself as the learner. Suppose you are watching a mystery movie on TV. Your OBA at any given time, is what you have gathered from the story up to that point in time. As you are watching, many scenes are presented to you; these are the MCFs. These scenes have different levels of importance and apparent relevance to the main theme of story. Some of these scenes are judged important and leave a sharply outstanding trace on your OBA. The trace left by others may not be as vivid. As you continue to watch, your understanding of the story gets updated. The updated OBA is the result of interaction between OBA and MCFs.

Notice that OBA remains contiguous at all times. The relevant MCFs get integrated into the OBA. Integration of MCFs into OBA account for the "ever changing, ever evolving, ever rearranging..." characteristic of the OBA. Also, the model has the capability to go back in time (dig into past experiences). Consider the case of watching mystery on TV. Often, as the story unfolds and the mystery is about to be solved, suddenly one or some of the past scenes (MCFs) which were not treated as important, get revived and play a significant role in the making of the conclusion and formation of the final experience. Thus, the OBA is seen to have the capability to mold old and new knowledge and come up with a more accurate account of the events.

Learning: Learning is defined as any change to OBA. Types of activities that the learning system (human brain) engages in from the time of first attention until the time of final conclusion in a learning episode are: attention, inquiry, closure, combination and knowledge creation. The learner pays attention to a certain process when, due to either internal activation or external excitation the system desires to learn about the process. Attention leads to inquiry; The system either mindfully or subconsciously becomes inquisitive about the process in question. Inquisitiveness leads to new data becoming available to the system (closure). This can be the result of externally triggered MCFs as well as OBA trying to revive the past knowledge. Once all types of data (OBA digging into the past, externally triggered MCFs, etc.) from different sources (different subsystems) about a certain event are available to the learner, the learner tries to make sense of all data (combination). If the process leads to or includes an experience worthy of remembering, it gets integrated into the OBA (knowledge creation).

4. Comparison

The comparison of some of the features of the theories that we reviewed is shown in the table below. Comparison of the theories that we discussed, reveals that biofunctionality fits best our originally intended learning model.

	Biofunctional	Associative and Connectionist
Paradigm	whole theme	piecemeal
learning Mechanism	double activity hypothesis (OBA and MCFs)	assembly line approach (three sequential phases) for IPT, single pattern processing for neural networks
Processing Capacity	the capacity of human brain	subject to limitations of explicit memory
Information Routing	information is available throughout the system	dependent on the existence of physical connection between brain and the organ
Types of Learning	allowing both intentional and incidental learning	intentional learning only (IPT)

5. Biofunctionality in Detail:

Biofunctionality and Remembering: The following analogy and examples demonstrate that the biofunctional model describes the human experience of remembering very closely.

Rubber Sheet Analogy: Recall OBA; "... an unbounded, ever present, ... contiguous **body of knowledge". This,** can be thought of as a rough surface with bumps, peaks and valleys, just like one of those computer generated terrain maps. Every feature (bump, peak or valley) of this surface corresponds to a certain experience (knowledge), of a network of interrelated (not entirely separable) experiences. Suppose this surface is made of rubber sheet. Now, consider this rough surfaced rubber sheet in a three dimensional Cartesian space. The vertical axis signifies the vividness of the memory. At the bottom of valleys, the awareness about a certain experience is almost non-existent (oblivion region). On top of the high peaks, the experience is vividly present in the memory (full-awareness region). Thus, remembering a certain event will be like pulling a valley or a low peak from oblivion or low awareness up to the full-awareness region, The higher the point comes up, more vividly it is remembered.

Remembering is not exact: Often, when trying to remember an old tune (a piece of music that you used to listen to some years ago), in addition to the main subject of recall (the piece of music), other memories of those days will come to life. You will probably remember the setup of your living quarters, taste of foods that you used to eat and other details. The surrounding items and details may not be recalled as vividly, nevertheless, most likely more memories will be revived than you originally intended to recall. This phenomenon can be explained by the rubber sheet remembering analogy. When you target a point on the sheet and pull it up, a cone like segment of the sheet will come up with it. The apex (main focus of attention) attains the highest degree of vividness. The surrounding material will become livelier than before but not as vivid as the main item to be remembered.

Repeated Remembering: When you try to remember an event which took place a couple of years ago that you have not thought of ever since, you might have to make some effort to remember it. However, if you make a second attempt a few days after the first, the second time around the recalling process will require much less effort. This can be explained by the rubber sheet remembering analogy. Forgetting an experience is equivalent to sinking down of its representative feature (peak, bump, .,.) to the oblivion region. This is like resetting of a mechanical system, which is usually a function of time. Rubber sheet too, can be considered to be slow in returning to its position before being disturbed. At the time of the second attempt, the newly risen peak has not totally retreated to its pre-remembering position. The longer the time lapse between the two remembering instances, the greater the effort for second remembering. This is of course in contrast with the explicit memory models. If you pick a book which has been sitting in a library shelf for the past few years, put it back and pick it up again the next day, the effort needed to withdraw the book, will be same in both cases.

Sources o Regulation." Learning and remembering in the biofunctional model are regulated by internal and external sources of control. The biofunctional model accommodates both intentional and incidental learning. To account for these two types of learning we define two internal sources of regulation; active and dynamic. The active (executive) type of control, regulates the intentional learning, while the dynamic (non-executive) control regulates the unintentional activities. An example of active control is the case when we deliberately initiate au action. A wide range of biological functions, such as heart beat, on the other hand are dynamically regulated.

A Human Experience Example: Let us review the first example of section 3.3. Consider yourself as the learner and your impressions of someone who you regularly interact with as the OBA. Any time that you notice the individual in question, or come in contact with him, your attention gets activated. If the individual utters some words towards you, naturally you will notice and make an attempt to hear him out. This accounts for the inquiry sub-function. The receiving of the message is the closure. Through the message the individual may alter your impression of him, remind you of a promise (cause you to remember and compare), or give you some information, These activities will all take place at the combination stage. If a conclusion is drawn or a piece of information worthy of future remembering is received, it w ill be integrated into knowledge base during the knowledge creation.

A More Technical Set Up: Consider a distributed system set-up composed of a robot (the host of learning), a number of users and resources and an information network. The robot's job is to fetch the resources requested by the users. Jobs have different priorities. Any time that a user desires a resource, it announces the request in the information network. Similarly, the resources will announce their being or not being available. All the announcements are atomic (they either reach all the players or are not received by anyone). The robot has some initial familiarity (OBA) with the set up of the environment in which it is operating. The attention sub-function is always on the look out for new announcements on the information network. When a user requests a service, the attention sub-function and subsequently the inquiry are activated. During closure the message is received by the robot and is interpreted. The robot inquires about the availability of requested resource and the priority status of the job. It may use auxiliary functions such as path finder or priority scheduling to construct the response (combination). After some operation time the robot may pick up certain patterns of job referrals (habits) which may help it schedule the jobs more efficiently or develop some navigational (obstacle avoidance) rules. The robot should also make note of permanent unavailability of a resource (the resource being used up). The knowledge creation sub-function will integrate such data (with long term usage potential) into the OBA.

6. Conclusions

The comparison between the reviewed theories and biofunctionality reveals that the biofunctional model accommodates the characteristics and requirements of our originally intended learning model. These requirements were: contiguous body of knowledge, universal information exchange and connection-independent execution environment. The model is particularly suitable to dealing with complicated situations with inter-dependent players, where the traditional single thread analysis may not yield the best result. We believe that the biofunctional model can successfully and accurately model the decision making aspects of a multi variable operational environment. We intend to employ the refined biofunctional learning and remembering model to a multi robot task scheduling and task performing scheme similar to the scenario discussed above.

7. Acknowledgment

This work is partially funded by a grant from Autonomous Control Engineering Center ACE-48146. The first and second authors wish to thank them for their financial support.

8. References:

1- Grossberg, Stephen, "The Attentive Brain", American Scientist, 1995, Volume 83, pp. 438-449.

2- Grossberg, Stephen, Carpenter, A. Gail, "Integrating and Neural Processing in a Self-Organizing Architecture", Artificial Intelligence and Neural Networks, 1994, pp. 387-421.

3- Iran-Nejad, Ali, Marsh, E. George, "The Figure and the Ground of Constructive Brain Functioning: Beyond Explicit Memory Processes", Educational Psychologist, 1992, 7(4), pp. 473-492.

4- Homaifar, Abdollah, Iran-Nejad, Ali, "Associative and Non-associative Theories of Distributed Learning and Remembering", S. J. Schmidt (Ed.), Frankfurt, Germany.

5- Iran-Nejad, Ali, "Associative and Nonassociative Schema Theories of Learning", Bulletin of the Psychonomic Society, 1989, 27(1), pp. 1-4,

REGIONAL GEOLOGIC MAPPING OF MESOZOIC REDBED SEQUENCES IN NORTHERN MEXICO UTILIZING LANDSAT THEMATIC IMAGES.

Claudio Bartolini
PACES
Department of Geological Sciences
The University of Texas at El Paso

STUDY AREA

The area of study for this investigation is comprised of two parts: a) the study area and the b) map area (Figure 1). The area of study is the area in northern and central mexico where stratigraphic sections of the Nazas Formation and the Huizachal Group were measured. The study area comprises the states of Coahuila, Durango, San Luis Potosi, Zacatecas, Nuevo Leon, and Tamaulipas.

The map area refers to the Caopas-Apizolaya Quadrangle in northern Zacatecas where approximately 3,000 km^2 were mapped utilizing Landsat thematic images. In addition, detailed stratigraphic studies, geochemistry, and geochronology of the Caopas, Rodeo, and Nazas Formation were also performed.

OBJECTIVES AND SIGNIFICANCE OF STUDY

Studies of Mesozoic redbed successions in north and central Mexico are very scarce, and confined primarily to the La Boca and La Joya Formations outcrops in the area of Ciudad Victoria in the state of Tamaulipas. Geologic mapping of red beds is sparse and mainly focused on the redbed outcrops of Tamaulipas. Studies of the redbed successions in the State of Nuevo Leon are also very limited.

The first objective of this investigation is to prepare geologic maps and structural sections of Triassic-Jurassic Nazas outcrops in the San Julian Anticlinorium, Zacatecas, using a Landsat imagery cover more than 3,000 square kilometers. This area is particularly critical because it encompasses the largest outcrops of the Mesozoic Nazas Formation in northern and central Mexico. Regional structural analysis, in particular, will document the structural and tectonic evolution of this region, which will be integrated to a more tectonic regional analysis. The new data will yield insights into the possible controls of basement tectonics on the development and configuration of the Jurassic-Cretaceous marine platform, and its subsequent evolution into the fold and thrust belt that now constitutes the Sierra Madre Oriental in northern and eastern Mexico.

The second goal of this study is to integrate the stratigraphy, depositional facies, paleocurrent data, and paleontology of the Triassic-Jurassic sequences to provide a regional paleogeographic overview. The history of the Mesozoic continental basins,

their original geometry and configuration in northern Mexico will be approached by integrating the stratigraphic, sedimentological, structural, geochemical, and isotopic data. The analysis will be complemented by the geochemical and isotopic data obtained from the volcanic rocks associated with the sedimentary sequences. This objective will be achieved through detailed description and measurement of the volcanic-sedimentary sections of the Nazas Formation in several basins located in the states of Durango, Coahuila, Zacatecas, San Luis Potosi, Nuevo Leon, and Chihuahua. Primary sedimentary structures and sedimentological characteristics will also be studied to interpret the regional depositional framework. Sandstone and volcanic samples collected from the measured sections will be prepared for petrographic study. Sandstone detrital modes will yield information about regional and temporal changes in sandstone provenance, possible sediment sources, volcaniclastic input, and interpretation of tectonic setting.

The third goal of this study is to constrain the age and affinity of volcanic rocks within the Nazas Formation. Isotopic ages (Ar-Ar and K-Ar) will be essential not only to correlate the stratigraphic sections, but also to constrain the timing of volcanic events during basin evolution. Selected volcanic samples also will be analyzed for their major and trace elements. These analyses will determine the volcanic lithologies present, and their tectonic affinity (i.e., whether the volcanic rocks are related to extension-related volcanism, or subduction-related volcanism). For example, the most distinctive chemical features of continental-margin volcanic suites are higher concentrations of K, Sr, Rb, Ba, Zr, Th and U, K/Rb, and Fe/Mg ratios, whereas the chemical composition of erupted magmas in continental rift zones is more complex, but generally contents higher concentrations of Mg, Al, Cr, Ni, Co, Nd, and Pb, and constant Nb/Zr and Hf/La ratios (Wilson, 1993).

The magmatic history and tectonic framework of the Nazas Formation will be important contributions to the geology of Mexico, especially if they are supported by geochronologic and geochemical data.

METHODOLOGY

Mapping.

Regional geologic maps of Mesozoic regions were produced from Sierra de San Julian, Sierra de Teyra, and Sierra de Candelaria, Zacatecas were produced at 1: 50, 000 scale. This area in northern Zacatecas is here considered the key map area for this investigation and was mapped in detail with the aid of Landsat imagery. The Landsat thematic mapper (TM) scene also covers a critical areas in northern and central Mexico, where the most extensive and best exposed outcrops of Mesozoic Nazas occur underneath the fold and thrust belt. The scenes were processed at the Jet Propulsion Laboratory in Pasadena, California. This digital data processing consists of creating false color composite images using TM bands 4 (Reflected near IR), 3 (visible red), and 7 (Reflected short wavelength H?). This *multispectral* combination *yields the* best enhancement of geologic features, based on preliminary image processing results that have already been completed. Subsequently, the TM data were superimposed on digital elevation models to obtain perspective views of selected areas within the San

Julian Anticlinorium in northern Zacatecas. The combined analysis of multispectral and topographic data provides a powerful tool for acquisition of stratigraphic structural information.

For the purpose of mapping, the TM image of northern Zacatecas was divided into subimages covering the areas of interest. Geological mapping was essentially done at a scale of 1:50,000, which is the same scale as Mexican topographic maps, whereas a regional structural interpretation of the map area along with a significant portion of the fold belt was done at a 1:250, 000 scale. The 1:50, 000 scale images are also appropriate for the geologic detail needed, and the nature of the outcrops. In addition, other subimages were printed at 1:100, 000 scale to interpret large-scale structures which are only partially seen on the 1:50,000 scale ones.

Stratigraphic Work.

The localities where known major redbed outcrops occur are shown in Figure 1. Measurement of stratigraphic sections with tape and Brunton was done in those localities where the deformation is not very complex, However, other sections were measured in intensely metamorphosed areas. Some large areas required of the measurement of two or more sections due to lateral stratigraphic variations within the area. Field methods included the construction and measurement of stratigraphic columns, identification and description of hand sample rock specimens, and the collection of sedimentary and volcanic samples for further study. Contacts within the redbed units, and also between older and younger formations were also described to understand the chronology of geologic events. Other data such as bedding characteristics, grain size, primary sedimentary structures, and paleocurrent measurements were also collected. Structural features such as folds, fault types, metamorphic fabrics and fractures were identified, described, measured, and mapped in the field. Simplified geologic maps of Mesozoic localities in other states where sections were measured were constructed from previous published data.

Laboratory work.

Thin sections of igneous and sedimentary samples collected from the measured sections were prepared, stained for feldspar recognition, and petrographically described.A total of 240 thin sections were studied under the microscope. A total of fifty one selected volcanic rocks (approximately 100 gr.) were pulverized at the crushing room in the UTEP Department of Geological Sciences. The samples were sent to XRAL laboratories in Ontario, Canada sent for standard major (SiO_2, TiO_2, Al2O_3, Fe2O_3, FeO, MnO, MgO, CaO, Na2O, K_2O, P_2O_5) and trace element (Rb, Sr, Ba, Zr, La, Ce, etc.) with 20 ppm detection limits.

Geochronology

The isotopic analysis (Ar-Ar) of 15 volcanic rock samples of the Nazas from the states of Durango, Zacatecas and San Luis Potosi were performed at the Laboratory of geochronology in the University of Houston. A total of four samples were dated using the K-Ar whole-rock method at the Geochron Laboratory. One sample was collected at the Caballeros Canyon, Tamaulipas, another at the San Marcos area, Nuevo Leon,

another one at Miquihuana, Tamaulipas, and the fourth one at the Cerro La Cruz (Aramberri), Nuevo Leon.

Geochemistry.
A total of 45 samples of volcanic rocks were collected from the Nazas Formation in the states of Durango, Coahuila, Zacatecas, and San Luis Potosi. Four more samples were collected in the Ciudad Victoria Canyons, Tamaulipas, and one in the San Marcos area, Nuevo Leon. The volcanic rocks selected for geochemical analyses are considered "relatively fresh" and non-foliated. Considering the Mesozoic magmatic history of Mexico, completely fresh rocks do not exist. Geochemical analyses were performed at XRAL Laboratory in Canada. Major element geochemistry is reported as SiO_2, Al_2O_3, CaO, MgO, Na_2O, K_2O, Fe_2O_3, MnO, TiO_2, P_2O_5, Cr_2O_3, and LOI, and trace element geochemistry is reported on Rb, Sr, Y, Zr, Nb, Ba. Geochemical analyses were done using the X-ray fluorescence method #I 02.

CONCLUSIONS

1) Late Triassic to Middle Jurassic volcanic-sedimentary sequences of the Nazas Formation were formed within a continental margin volcanic arc that extended northwest, across Mexico. This volcanic arc is the southern continuation of the magmatic arc of western North America.
2) Triassic to Jurassic redbed strata of the Huizachal Group accumulated in rift basins which are related to the rifting and opening of the Gulf of Mexico basin.
3) Geologic mapping with Landsat images allowed the recognition of outcrops of the Nazas Formation within the cores of breached anticlines along the Mexican fold belt. Further detailed mapping of these rocks defined a more precise lithologic division of this formation.
4) The characterization of structural styles developed by the lower Mesozoic continental strata and the overlying carbonate platform was only possible with the construction of multiple perspective views.

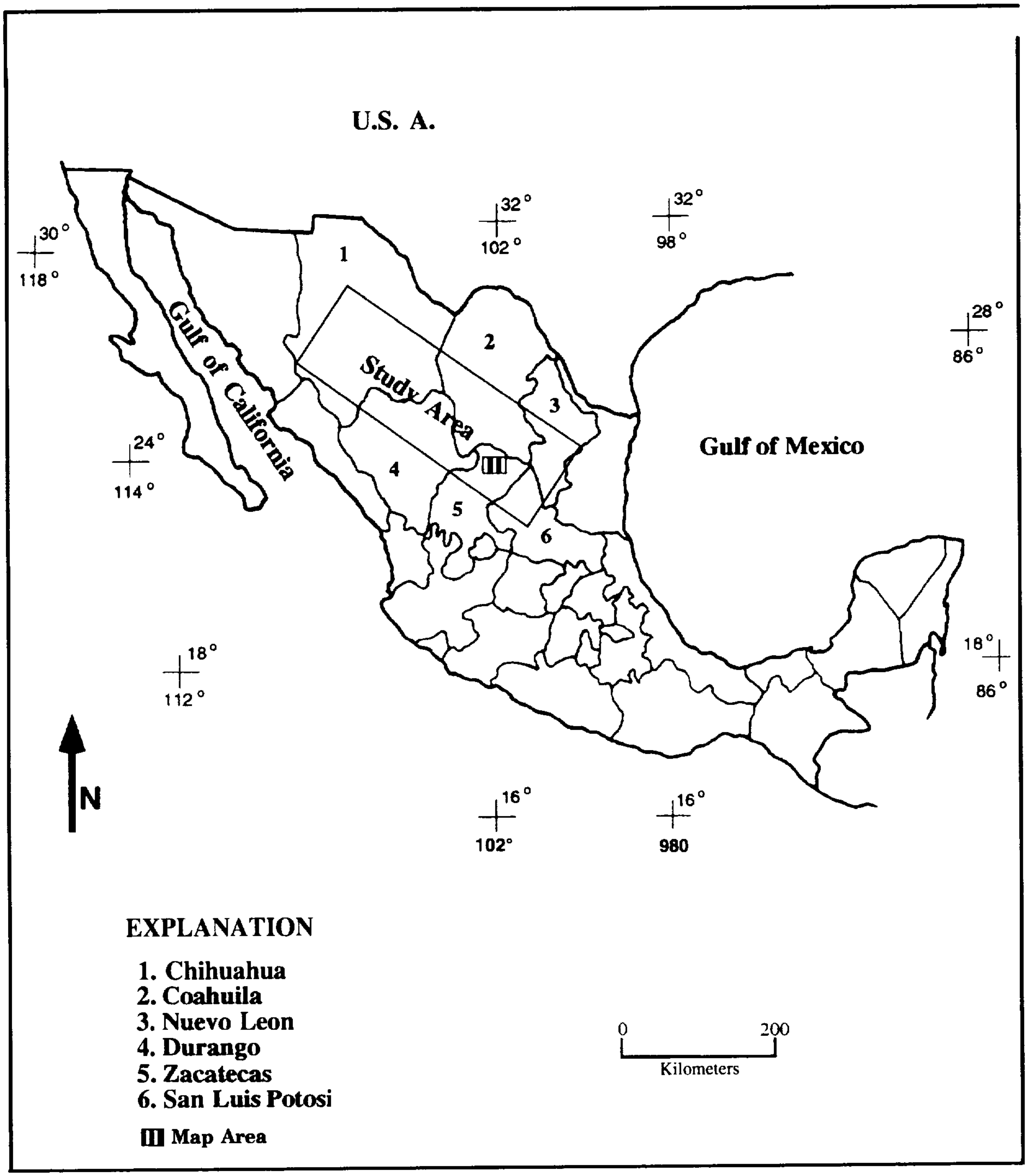
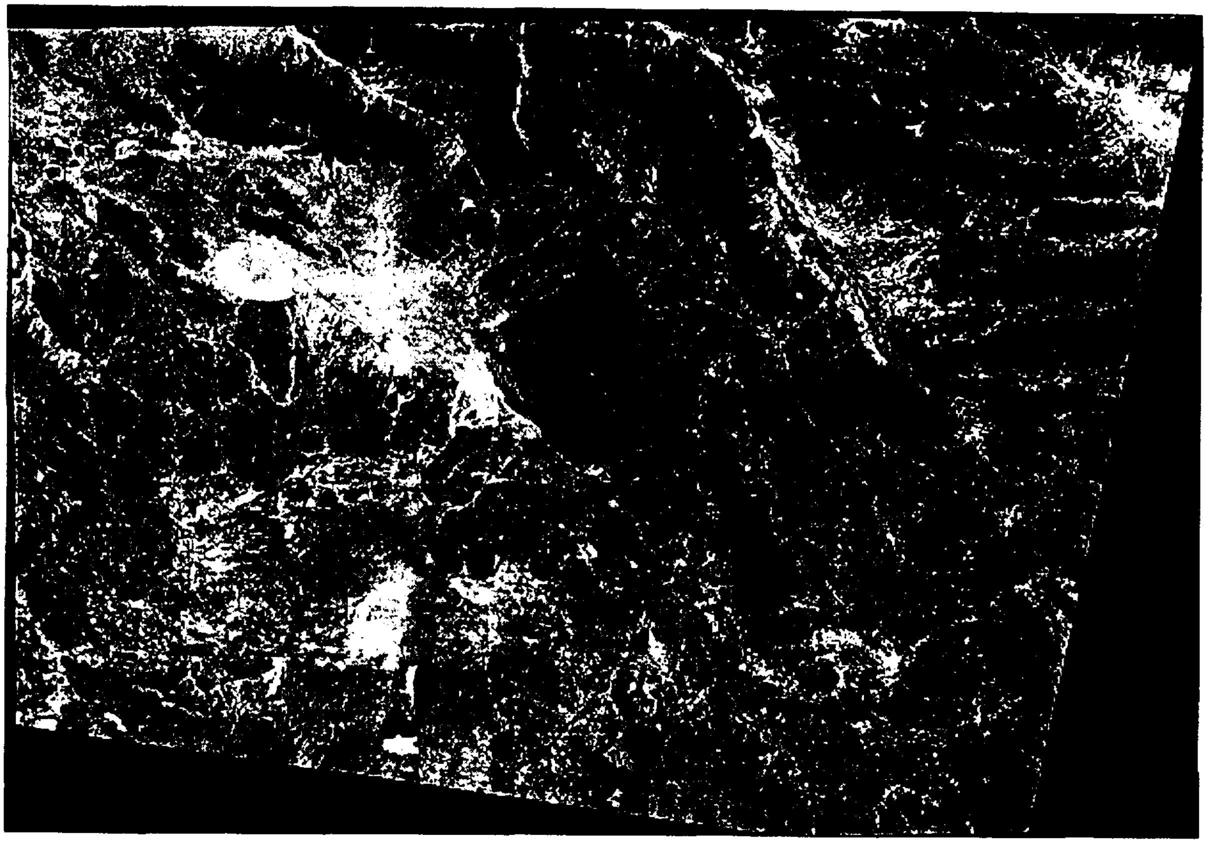


Figure 1. Approximate loacation of study and map areas in northern and central Mexico.

San Julian Anticlinorium

Thematic Mapper band 5



This image is the fifth channel of the Thematic Mapper scanner mounted on the Landsat 4 and 5 satellite. Its wavelength is in the reflected infrared band of the electromagnetic spectrum and is outside the range of human vision. Its specific wavelength is 1.55μ to 1.75μ. This spectral range is useful for determining moisture content of soil and vegetation. This band penetrates thin clouds and is provides good contrast between vegetation types. This image has been contrast enhanced using a gaussian stretch with a sigma value of 2.5 and a light high pass filter applied to bring out detail. The raw un-enhanced version of this file is available.

ULTRASOUND ALGORITHM DERIVATION FOR SOIL MOISTURE CONTENT ESTIMATION

Belisle, W. R., R. Metzl, J. Choi, M. D. Aggarwal, and T. Coleman
Center for Hydrology, Soil Climatology, and Remote Sensing
Alabama A&M University, Normal, Alabama 35762

ABSTRACT

Soil moisture content can be estimated by evaluating the velocity at which sound waves travel through a known volume of solid material. This research involved the development of three soil algorithms relating the moisture content to the velocity at which sound waves moved through dry and moist media. Pressure and shear wave propagation equations were used in conjunction with soil property descriptions to derive algorithms appropriate for describing the effects of moisture content variation on the velocity of sound waves in soils with and without complete soil pore water volumes, An elementary algorithm was used to estimate soil moisture contents ranging from 0.08 g/g to 0.5 g/g from sound wave velocities ranging from 526 m/s to 664 m/s. Secondary algorithms were also used to estimate soil moisture content from sound wave velocities through soils with pores that were filled predominantly with air or water.

INTRODUCTION

Ultrasound methods involve the mechanical vibration and propagation of waves above about 20,000 cycles/s through various materials (Dull et al., 1964), The velocity of these propagating waves is affected by the nature of the material through which it is passing. Curtis (1982), Szilard (1982), and Kinsler et al. (1982) describe several equations and input parameters used in estimating the velocity of sound waves traveling through different media. Sound waves travel through air and water at velocities of approximately 330 m/s and 1660 m/s, respectively. These waves travel through solid materials such as aluminum and steel at velocities of 6,300 m/s and 6,100 m/s, respectively. They also travel through porous materials such as concrete, ice, and cork at velocities of 3,100 m/s, 3,200 rids, and 500 m/s, respectively (Kinsler et al., 1982). Kinsler et al. (1982) also reported a number of other solid material properties associated with sound wave velocties including density (kg/m^3), Young's Modulus (Pa), Shear Modulus (Pa), Adiabatic Bulk Modulus (Pa), Poisson's Ratio, and Characteristic Impedance (Pa*s/m).

In-situ soils have some properties that are similar and different than those of many materials classically evaluated using ultrasonic sound wave velocity measurements. Soils are quite heterogeneous and the classically evaluated solid materials are homogeneous (Brady, 1990; Szilard, 1982). Soils can be compacted to a greater degree than materials such as steel and iron and have porosities around the range of materials such as cork and oak wood. Moisture content variation in soils occurs differently than in solid materials such as steel and iron, or glass and quartz. Texture, organic matter content, bulk density and porosity are major contributors to the soil water holding capacity (Brady, 1990). Additional contributions may be made by the presence of soil microbes, rocks, and other chemical and physical parameters. The purpose of this study is to perform initial investigations of the derivation of elementary and secondary ultrasound algorithms useful in evaluating soil moisture content variation. The study will attempt to incorporate soil moisture content, porosity, organic matter content, bulk density, and texture into existing equations that are used to estimate the velocity at which ultrasonic waves propagate through soils with pores that are filled predominantly with air or water.

ALGORITHMS

Sound waves moving through media with no boundary effects can be described by the sound velocity, c, which can be described by shear and pressure waves (Szilard, 1982).

$$c_{shear} = \sqrt{\frac{E}{\rho}\frac{1}{2(1+\sigma)}} = \sqrt{\frac{G}{\rho}} \qquad (1)$$

and

$$C_{pressure} = \sqrt{\frac{E}{\rho}\frac{1-\sigma}{(1+\sigma)(1-2\sigma)}} \tag{2}$$

where velocity is determined by the density, p, the modulus or elasticity, E, and Poisson's ratio, σ, and G is the modulus of rigidity (also called the shear modulus or torsional modulus). The density of a given volume of solid material is described similarly to the density of a given volume of soil, ρ_B; density of soils is equal to the mass of the dry soil, M_s, divided by the entire or bulk volume occupied by the material as shown in Equation (3)

$$\rho_B = \frac{M_S}{V_{TOTAL}} \tag{3}$$

where M_S is the soil mass and V_{TOTAL} is the total volume of a given soil sample. The total velocity, c_{TOTAL}, of sound waves through a medium may be described as

$$C_{TOTAL} = C_{press} + C_{shear} \tag{4}$$

Therefore,

$$C_{TOTAL} = \sqrt{\frac{E}{\rho}\frac{1-\sigma}{(1+\sigma)(1-2\sigma)}} + \sqrt{\frac{G}{\rho}} \tag{5}$$

Equation (5) suggests that the velocity of sound waves moving through solid media decreases with positive increases in p.

Soil moisture content, θ, is defined as

$$\theta = \frac{M_w}{M_s} \quad and \quad M_s = \frac{M_w}{e} \tag{6}$$

Substituting Equation (6) into Equation (3) results in

$$\rho_B = \frac{M_w}{\theta_m V_{TOTAL}} \tag{7}$$

Substituting Equation (7) into Equation (5) results in

$$C_{TOTAL} = \sqrt{\frac{E\theta_m V_{TOTAL}}{M_w}\frac{1-\sigma}{(1+\sigma)(1-2\sigma)}} + \sqrt{\frac{G\theta_m V_{TOTAL}}{M_w}} \tag{8}$$

C_{TOTAL} can be further defined as

$$C_T = \sqrt{\frac{E\theta_m}{M_w}\alpha} + \sqrt{\theta_m \beta} \tag{8a}$$

where

$$\alpha = V_{TOTAL}\frac{1-\sigma}{(1-\ \sigma)(1\ -20)} \tag{8b}$$

and

$$\beta = \frac{G\, V_{TOTAL}}{M_w} \tag{SC}$$

Since

$$\theta_v = \rho_B \theta_m \tag{9}$$

Equation (9) can be substituted into Equation (8) and results in

$$c_{TOTAL} = \sqrt{\frac{E \frac{\theta_v}{\rho_B} V_{TOTAL}}{M_w} \frac{1-\sigma}{(1+\sigma)(1-2\sigma)}} + \sqrt{\frac{G \frac{\theta_v}{\rho_B} V_{TOTAL}}{M_w}} \tag{lo}$$

Equations (8) and (10) are the elementary algorithms (EI and EII) used for estimating mass and volumetric soil moisture content, respectively, from sound wave velocity through soils. The effects of porosity on c_{TOTAL} *can* be described using a modification of either of the elementary algorithms as shown in Equation (11) and (12).

$$c_{TOTAL} = \sqrt{\frac{E\theta_m \frac{V_{SOLID}}{1-¢}}{M_w} \frac{1-\sigma}{(1+\sigma)(1-2\sigma)}} + \sqrt{\frac{G\theta_m \frac{V_{SOLID}}{1-\phi}}{M_w}} \tag{11}$$

$$c_{TOTAL} = \sqrt{\frac{E \frac{\theta_v}{\rho_B} \frac{V_{SOLID}}{1-\phi}}{M_w} \frac{1-\sigma}{(1+\sigma)(1-2\sigma)}} + \sqrt{\frac{G \frac{\theta_v}{\rho_B} \frac{V_{SOLID}}{1-\phi}}{M_w}} \tag{12}$$

where V_{SOLID} is the volume of solid portion of the soil system. Organic matter and texture effects can also be evaluated through the measurement and evaluation of the volume of soil organic and mineral fractions within V_{SOLID}.

Szilard (1982) further describes the velocity of ultrasound waves through air and water as

$$c_{AIR} = \sqrt{\frac{\gamma P_o}{\rho_o}} \tag{13}$$

and

$$c_{FLUID} \sqrt{\frac{k}{p}} - \tag{14}$$

where c_{FLUID} *is* the speed of sound waves through air, y is the specific heat ratio, P_o is the static pressure, ρ_o is the static density, c_{FLUID} *is* the speed of sound waves through water, k is the bulk stiffness modulus, and p is the density. Soil particles are considered to be in air until the water in the pore space flows or has similar properties as a complete pore volume of water. Equation (7) can therefore be substituted into Equations (13) and (14) to result in

$$c_{AIR} = \sqrt{\frac{\gamma P_o \theta_m V_{TOTAL}}{M_w}} \tag{15}$$

and

$$C_{FLUID} = \sqrt{\frac{k\theta_m V_{TOTAL}}{M_w}} \tag{16}$$

Equations (15) and (16) are the secondary algorithms (SI and SII) used for evaluating c through porous media in air and water, respectively. C_{AIR} and C_{FLUID} can be further defined as

$$C_{AIR} = \sqrt{\theta_m A} \tag{15a}$$

and

$$C_{FLUID} = \sqrt{\theta_m B} \tag{16a}$$

where

$$A = \frac{\gamma P_o V_{TOTAL}}{M_w} \tag{15b}$$

and

$$B = \frac{k V_{TOTAL}}{M_w} \tag{16b}$$

RESULTS AND DISCUSSION

The EI elementary algorithm (Equations 8a, b, and c) for estimating sound wave velocity from mass soil moisture content was used for the algorithm simulation. 100 gm of a well granulated soil and the preliminary data from Choi et al. (1996) were used as input data. The value of a was assigned a value of 0.3 and the shear propagation through the media was assumed to be negligible. The mass moisture content for the simulation ranged from 0.08 (wilting point) to 0.5 (saturation) and the numerical value for the modulus of elasticity, E, ranged from 9.25E5 (moist) to 1.47E6 (dry). Figure 1 shows the results of plotting θ_m versus c_{TOTAL} for the algorithm simulation. The velocity of the sound waves decreased from 664 m/s to 527 m/s with a slope of -312 as the soil moisture content increased from 0.08 g/g to 0.5 g/g.

Soil moisture contents ranging from 0.08 g/g (wilting point) to 0.5 g/g (saturation) were estimated from the velocity of sound waves moving through the soil media using the secondary algorithms SI and SII (Equations 15a and b and 16a and b). The results of plotting mass moisture content, θ_m, versus sound wave velocity in soils with little water in the soil pores, C_{AIR}, and in soils with predominantly water-filled pores, C_{FLUID}, are shown in Figure 2. The SI algorithm estimates soil moisture content ranging approximately zero to 0.3 g/g and the SII algorithm estimates soil moisture content ranging from approximate y 0.3 to that of saturation.

CONCLUSION

The ultrasound algorithms developed for estimating soil moisture content from sound wave velocities are viable tools which can be used in this regard. The EI and EII elementary algorithms are capable of estimating soil mass and volumetric moisture contents as well as other soil properties such as porosity, organic matter, and texture. The SI and SII secondary algorithms facilitate the estimation of soil moisture content from sound wave velocities at moisture contents ranging from approximate y zero to that of saturation. The major considerations and limitations associated with the algorithms include the fact that the equations described by Szilard were developed for solid, homogeneous materials as opposed to soils which are porous, heterogeneous materials. Material lattice description and consistency and the modulus of elasticity are parameter considerations that vary greatly from one material to the next. The described algorithms suggest a decrease in the porous material elasticity with increased moisture content.

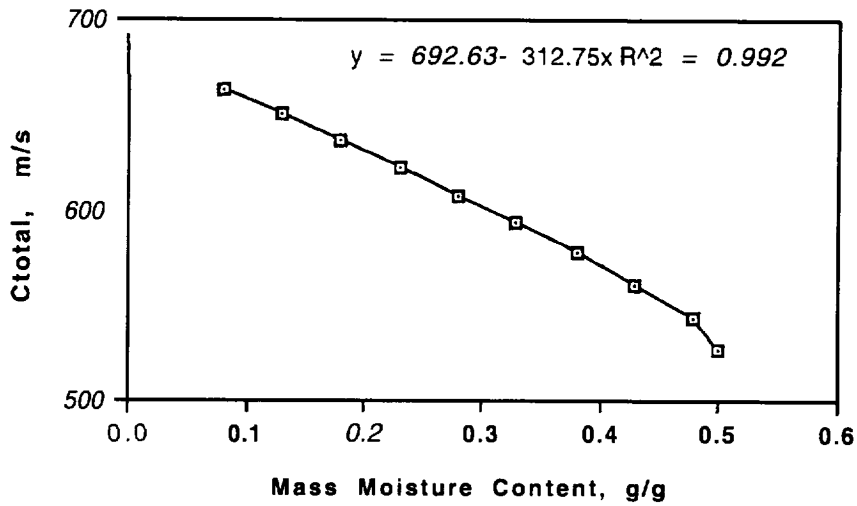


Figure 1. Mass moisture content versus soundwave velocity using the elementary algorithm, EI.

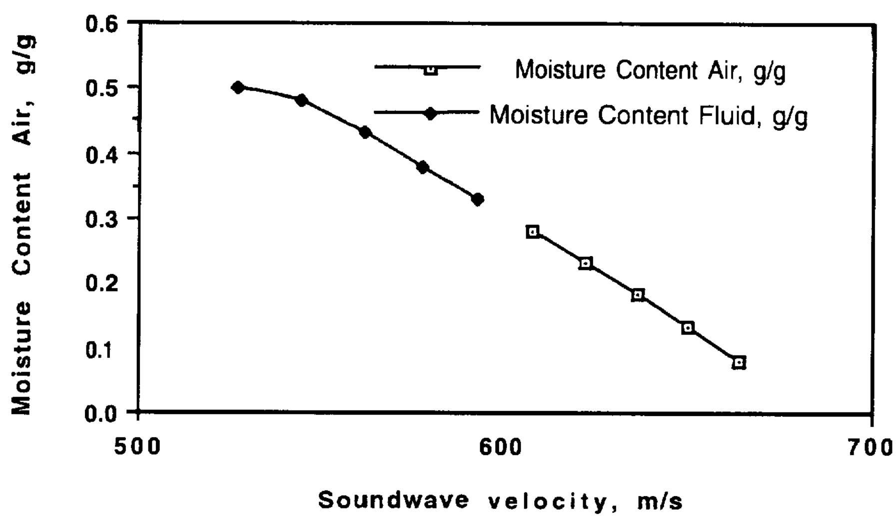


Figure 2. Soundwave velocity through soils containing predominantly air and water using the secondary algorithms, SI and SII.

ACKNOWLEDGMENTS

Acknowledgment is extended to the Center for Hydrology, Soil Climatology, and Remote Sensing (HSCaRS) support staff. Contribution from HSCaRS, the Plant, Soil, and Animal Sciences, and the Alabama Agricultural Experiment Station, Alabama A&M University, Normal, AL 35762, Journal No. 342. This work was supported by Grant No. NCCW-0084 from the National Aeronautics and Space Administration (NASA), Washington, DC. Any use of trade, product or firm names is for descriptive purposes only and does not imply endorsement by the U. S. Government.

REFERENCES

Brady, N. C. 1990. The Nature and Properties of Soils, 10th Edition, Macmillan Publishing Company, New York.

Choi, J., R. Metzl, M. D. Aggarwal, W. Belisle, and T. Coleman. 1996. Use of ultrasonic technology for soil moisture measurement, NASA University Research Centers Technical Conference on Education, Aeronautics, Space, Autonomy, Earth, and Environment Proceedings, Long Beach, California, February 16-19.

Curtis, G. J. 1982. Ultrasonic Testing, Non-destructive testing of adhesively bonded structures with acoustic methods, John Wiley & Sons, New York, Chapter 13.

Dull, C. E., H. C. Metcalfe, and J. E. Williams. 1964. Modern Physics, Holt, Rinehart, and Winston, Inc., New York, p. 298-313.

Kinser, L. E., A. R. Frey, A. B. Coppens, and J. V. Sanders. 1982. Fundamentals of Acoustics, 3rd Edition, John Wiley & Sons, New York, Appendix 10.

Szilard, J. 1982. Ultrasonic Testing, Non-conventional testing techniques, John Wiley & S ens, Chichester, p. 1-6.

URC97018

5/3/25

The Ulam Index: Methods of Theoretical Computer Science Help in Identifying Chemical Substances

Adriana Beltran[1,2] and James Salvador[2]

[1] Center for Theoretical Research
and its Applications in Computer Science (TRACS),
Department of Computer Science, and
[2]Department of Chemistry
University of Texas at El Paso
El Paso, TX 79968
emails [1]abeltran@cs.utep.edu and
[2]james@salvador.chemistry.utep.edu.

Abstract

in this paper, we show how methods developed for solving a theoretical computer problem of graph isomorphism are used in structural chemistry. We also discuss potential applications of these methods to exobiology: the search for life outside Earth.

1 Identification of chemical substances: why we need it, why it is difficult, and what we are going to do about it

Identification of chemical substances can be reduced to a graph isomorphism problem (well known in theoretical computer science). One of the main problems of chemistry is identification of chemical substances.

In non-organic and organic chemistry, there exist experimental techniques that enable us to describe a *graph* structure of the unknown substance, i.e., to describe which atoms it consists of, and which of these atoms are connected by chemical bounds. In order to identify this substance, we must compare it with graphs that describe known substances.

In mathematical terms, we need to check whether an (experimentally obtained) graph is isomorphic to one of the graphs that describe known substances.

Graph isomorphism problem is known to be hard. Unfortunately, the general graph isomorphism problem is known to be hard to solve.

For some substances, different nodes correspond to different types of atoms; in this case, it is relatively easy to check whether a given molecule coincides with this substance, because we can simply identify each atom with a similar atom in the standard substance and then check whether all connections are as in the standard model.

For many other substances. however, atoms of the same type occur in different places of the structure *in* different roles; examples of such substances are organic substances and fullerenes. For these substances, we have to actually solve the difficult graph isomorphism problem.

How to solve this difficult problem: the main idea. One way of solving this problem is based on the following idea:

- It is known that to every graph, we can assign a polynomial or several polynomials that uniquely determine this graph (i.e., the two tuples of polynomials coincide iff the graphs are isomorphic).
- Thus, to check whether the two graphs are isomorphic, we can compare the coefficients of the corresponding polynomials.

These methods are widely used in structural chemistry; see, e.g., [8, 1, 7, 9, 10, 2, 3].

We can further compress these polynomials into *numbers* (called *indices*) that also give complete information about the graph [1 I], and compare only these numbers.

A word of warning: index methods are only heuristic. The resulting methods are, of course, only **heuristic** method, because sometimes, due to computer inaccuracy, non-isomorphic substances get erroneously identified.

How frequent are the errors? Since the index methods are purely heuristic, it is important to check how frequently the methods err.

Our numerical experiments show that these errors are extremely rare (and that, therefore, this method works really well) [4]: among all the generated graphs, only 10-5 of them got mis-identified.

2 Some technical details

The index that we use. in this work, we use an index called *Ulam index* because it originated with the ideas presented by S. Ulam in [12].

The Ulam Index is defined (and calculated) as the result of substituting the properly coded structural information of *Ulam Subgraphs* (defined in [12]) into the ***matching polynomial*** of a graph (for a definition of a matching polynomial, see, e.g., [7]; the matching polynomial is a unique and invariant representation of a graph).

We want to substitute some values into the matching polynomial and get an index. To prevent two graphs from having the same index, we differentiate between the variables that correspond to different vertices by counting the number of times that each variable representing a vertex appears in the matching polynomial. This idea is similar to the one used in the definition of the Hosoya's Z index of the graph with that vertex deleted [9].

So, the first natural idea is to use these numbers of times as values of the variables that are substituted into the matching polynomial. This first idea leads to a good index, but, unfortunately, the resulting numbers are too large and cannot be easily represented in the computer.

In order to avoid this problem, before we substitute the weights, we *normalize* them bus dividing each weight by the Hosoya's Z Index of the whole graph (i.e., by the total number of terms **in** the matching polynomial). The result of substituting these normalized weights is what we call an Ulam index.

We have a program that computes the Ulam index. We have developed a computer program named GRADE (Graph Recognition Algorithm Developed for Education) that computes the Ulam index. This program is used, in particular, to tutor and test students in chemical nomenclature.

Ulam index is highly discriminating. The *Ulam Index is* a number that uniquely represents a planar graph. This index is ***highly discriminating*** in the sense that usually, non-isomorphic graphs have drastically different values of the Ulam index and therefore, even if we perform computations on real-life computers with computational inaccuracies, the resulting indices typically remain different.

In particular, as our computer experiments show, the Ulam Index differentiates all trees up to 20 vertices (there are 1,346,024 of them) and all graphs up to nine vertices (there are 274,668 of them).

3 Possible applications to space exploration

One of the major tasks for the past and future space missions to planets and other celestial bodes (such as comets and asteroids) has been to look for life or at least for traces of the former life (see, e.g., [5, 6]). This is especially important now, when traces of life has been found in meteorites coming from Mars.

Automatic robotic missions must be able to analyze the substances that they find on the other planets and identify them.

For this identification, graph isomororphisms algorithms can be of great help.

Acknowledgments. The authors would like to acknowledge the financial support of NSF grant No. CHE-9414968 and of the NASA PACES Center. One of the authors (A. B.) is thankful to Drs. Ann Gates, Vladik Kreinovich, Luc Longpré, and Scott Starks for their valuable **comments and support.**

References

[1] J. Aihara and H. Hosoya, *Bull. Chem.* Sot. *Japan*, 1988, **Vol.** 61, pp. 2657 ff.

[2] R. A. Beezer and E. J. Farrell, "The matching polynomial of a regular graph", *Discrete Mathematics*, 1995, Vol. 137. No. 1, pp. 7-18.

[3] R. A. Beezer, E. J. Farrell, J. Riegsecker, and B. Smith, "Graphs with a minimum number of pairs of independent edges I: Matching polynomials", ***Bulletin of the Institute of Combinatorics and Its Applications, 1996*** (to appear).

[4] A. Beltran and J. M. Salvador, "The Ulam index", ***Abstracts of the Second SC'- COSMIC*** Conference *in Computational Sciences, **October 25-27, El Paso, TX,*** Rice University Center for Research on Parallel Computations and University of Texas at El Paso, 1996, p. 6.

[5] K. Biemann, J. Ore, I?. Toulmin, L. E. Orgel, A. O. Nier, D. M. Anderson. P. G. Simmonds, D. Flory, A. V. Diaz, D. R. Rushneck, J. E. Biller, and A. L. Lafleur, "The search for organic substances and inorganic volatile compounds in the surface of Mars", *J. Geophys.* **lies.,** 1977, Vol. 82, **pp.** 4641-4658.

[6] R. R. Christensen, D. L. Anderson, S. C. Chase, R. N. Clark, H. H. Kieffer, M. C. Malin, J. C. Pearl, J. Carpenter, N. Bandeira, F. G. Brown, and S. Silverman, "Thermal emission spectrometer experiment: The Mars Observer mission", *J. Geophys. Res.*, 1992, Vol. 97, pp. 7719-7734.

[7] I. M. Gutman and V. R. Rosenfeld, "A Novel Approach to Graph Polynomials", ***MATCH,*** 1989, **Vol. 24,** pp. 191-ff.

[8] H. Hosoya, *Comp. **Math. Appls., 1986, Vol.*** 12B, pp. 271 ff.

[9] H. Hosoya and K. Balasubramanian, "Computational Algorithms for Matching Polynomials of Graphs from the Characteristic Polynomials of Edge-Weighted Graphs", ***Journal of Computational Chemistry, 1989,*** Vol. **10,** No.5,pp. 698-710.

[10] M. Randic, H. Hosoya, and O. E. Polansky, "On the Construction of the Matching Polynomial for Unbranched Catacondensed Benzenoids", ***Journal of Computational*** *Chemistry, **1989,*** Vol. 10, No. 5, pp. 683--697.

[11] J. M. Salvador, "Topological indices and polynomials: the partial derivatives", ***Abstracts*** *of the 5th* ***International Conference on Mathematical and Computational Chemistry, May 17-21, 1993,*** Kansas City, Missouri, p. 154.

[12] S. M. Ulam, ***A Collection of Mathematical Problems,*** John Wiley and Sons. New York, 1960.

Page intentionally left blank

5/19/82

URC97019

Managing Data in a GIS Environment

Maria Beltran and Haris Yiasemis
Department of Computer Science
The University of Texas at El Paso
El Paso, TX 79968
{ mbeltran, haris }@cs.utep.edu

1 Introduction

A Geographic Information System (GIS) is a computer-based system that enables capture, modeling, manipulation, retrieval, analysis and presentation of geographically referenced data. A GIS operates in a dynamic environment of spatial and temporal information. This information is held in a database like any other information system, but performance is more of an issue for a geographic database than a tradition] database due to the nature of the data.

What distinguishes a GIS from other information systems is the spatial and temporal dimensions of the data [5] and the volume of data (several gigabytes). Most traditional information systems are usually based around tables ant] textual reports, whereas GIS requires the use of cartographic forms and other visualization techniques. Much of the data can be represented using computer graphics, but a GIS is not a graphics database. A graphical system is concerned with the manipulation and presentation of graphical objects whereas a GIS handles geographic objects that have not only spatial dimensions but non-visual, i e., attribute and components. Furthermore, the nature of the data on which a GIS operates makes the traditional relational database approach inadequate for retrieving data and answering queries that reference spatial data [.5].

The purpose of this paper is to describe the efficiency issues behind storage and retrieval of data within a GIS database. Section 2 gives a general background on GIS, and describes the issues involved in custom vs. commercial and hybrid vs. integrated geographic information systems. Section 3 describes the efficiency issues concerning the management of data within a GIS environment. The paper ends with a summary of the main concerns of this paper.

2 Geographic Information Systems (GIS)

2.1 Data

GIS data is handled in a single database or a collection of databases [4]. Like any other database, it must be secure, reliable, and consistent. Some special characteristics of GIS databases include:

- Spatial Data. Spatial Data is traditionally divided into two classes, raster and vector [1]. Raster data is structured as a two-dimensional array of cells or pixels. Each cell in a raster is addressed by its position in the array. A point may be represented by a single cell, and a connected area by a collection of contiguous cells. When a raster is stored with no compression, it can be extremely inefficient in terms of storage. The other type of data is vector data. A vector is a finite straight line segment defined by its end-points, and the locations of the end-points are given with respect to some coordinates of the plane.

- Data capture. Data capture involves two requirements. The first requirement is to provide the physical devices for capturing data external to the system and inputting it to the database. The second is to provide software for converting data to structures that are compatible with the data model of the database and checking for integrity of data before entry into the system. Geographic databases have a wider variety of sources and types of data than traditional databases. The main problem here is to get the data into a format that is acceptable by a particular GIS. Primary input devices used by GIS in addition to keyboard and voice recognition systems are as follows:

- *Remote* sensing captures data by means of sensors on a satellite that provide measurements of reflectance or images of portions of the earth. The data is usually raster in structure.
- *Global* Positioning *Systems (GPS)* allow the capture of terrestrial position and vehicle tracking, using a network of navigation satellites. Data is captured as a set of point position readings and are in vector format.

Secondary data capture usual] y is from paper-based maps, The following devices are used:

- Scanners convert an analog data source (e. g., a map) into a digital dataset in raster format.
- *Digitizers* convert an analog data source into a digital dataset in vector format.

Currently there is interest in new methods for raster-to-vector and vector-to-raster conversion.

- Data **retrieval**. Most interactions with a database are attempts to retrieve data [5], [4], [3]. A GIS allows real spatial processing to take place. Examples of spatial queries are:

  - What is at a particular location? This may be clone by clicking the mouse at a particular location on the screen or by giving coordinates.
  - What locations satisfy these requirements? For example, find the names of all land areas that satisfy the following requirements : 1) less than average price for lan d, and 2) within 15 minutes drive of 110.

  Performance is a bigger problem for a geographic database than a general-purpose database because of the volume of data. Also, the nature of the data is often hierarchical (e.g., a point is a part of an arc and an arc is a part of a polygon) and this creates difficulties for traditional database approaches. Special storage structures and access methods are required.

- Data **presentation.** Tradition! databases provide output in the form of text usually in tabular form. They also may generate reports with charts and other graphical displays. A GIS requires a more sophisticated presentation of results, which might be multi d dimensional [5],[2]. output might be in the form of maps and other sophisticated forms. Graphics and visualization tools are a key component of a GIS and include tools for the creation, storage, and manipulation of models and images of objects. Graphic images are essential to a GIS, but these images also need a huge amount of space for storage. Multimedia computing presents new opportunities for GIS, but also creates new problems because of space issues. Storage and compression of such d ata is a major research area [5].

- Data **distribution**. The trend in recent technology is to move from centralization towards a distributed computer system [3] in which machines communicate through a network. As a consequence, data and database management systems (DBMS) are distributed. through the network. In this way, the reliability of the system is achieved because failure at one site will not mean failure for the whole system. Furthermore, distributed data may be natural and appropriate for a GIS because particular data may be associated with a particular site, e.g., details of local weather conditions may be better held at a local site where local control and integrity checks [3] may be maintained.

2,2 Analytical **processing**

Analytical processing is one of the major requirements for GIS [5],[6]. Some of the requirements include the following:

- Geometric/topological analysis: Most geographically- referenced objects have geometric or topological properties. Topological operations include adjacency and connectivity relationships. Geometric analysis would involve locating a spot in a region considering the distance from a place of reference to the spot.

- **Terrain** and **field** analysis: Terrain analysis is usually based upon datasets giving topographical elevations of point locations, i.e., degree and direction of slope. This would include analyzing the visibility between locations. Spatial fields arc variations of attributes over a region, or the topographical elevation over an area. Fields may be scalar (variations of a scalar quantity represented as a surface) or vector (variations of a vector quantity like wind-velocity). Field operations include slope analysis, view-shed analysis and path finding.

- Network **analysis**: A network is a configuration of connections between nodes. Application of network analysis in GIS may be found in many areas from transportation networks to utilities. Network operations include connectivity analysis, path finding and flow analysis, An example would be to provide a route in order to visit all attractions of a tourist region minimizing the time.

2.3 Custom vs. Commercial GIS

Many commercial GIS packages exist today. Some of the commercial packages, such as ARC/INFO, provide a programming language and interfaces that allow the user, for instance, to access an external relational database package giving the user some flexibility in organizing the system. An organization may decide to develop and program a GIS from scratch to meet a special need (called a custom *GIS* in this paper) or to create application programs that interact with commercial packages.

Due to the large amount of data, and to the nature of spatial data, a custom GIS sometimes is preferred when fast access of data is required. The data in such cases is stored in an application-dependent way to provide this capability. The disadvantage in this approach is that it has only limited query capabilities because there is no connection with an external relational database package. Also, most of the query operations are hard-coded, sacrificing generality and flexibility for better query performance.

2.4 Hybrid vs. Integrated GIS

GIS can be categorized into two groups according to their general architecture: hybrid and integrated [5],[2]. Suppose that we want to keep spatial and non-spatial data about. a particular piece of land. Usually spatial data describes location in two or more dimensions, For example, the geometry of the land with topological relationships might be kept in a map. The name, address, and owner information constitute the non-spatial data. These non-spatial data might he kept in a relational database. On the other hand, the spatially referenced data are not immediately compatible with a relational database and must be stored in a proprietary database. This is because in order to store spatial data in a relational database, each dimension must occupy a separate column, making spatial queries very time consuming.

The primary reason for having a hybrid architecture is the distinction between spatial and non-spatial data. Usually in hybrid systems (e.g. ARC/INFO) spatial data is stored in a set of system files, and non-spatial(attribute) data is stored in a relational database. In such systems. one part forms the graphics and spatial data engine, and the other part, handles the non-spatial data in tile database. The advantage of this approach is that the search performance problem is minimized, but of course with the loss of generality. This approach has the disadvantage that, the spatial data are handled outside the database and cannot take advantage of the capabilities of relational database technology such as integrity, security and reliability. Furthermore, due to the proprietary database in this approach, the exchange of data between different databases is complicated, if not impossible.

The idea behind an integrated architecture is to manage the data in a single database. The spatial data is handled in the same way as non-spatial. The problem with such a solution is that performance on retrieval of spatial data is poor due to the large number of relational operations that, are required to reconstruct the spatial objects. Keeping everything in a traditional relational database solves some problems but creates others. This might explain why most GIS systems are based on the hybrid architecture.

3 Data Management in GIS

'the large volume of GIS data makes the traditional relational database approach inadequate for retrieving and managing data. Spatial and temporal data with possibly two or more dimensions must also be referenced making data management more complicated than a traditional information system.

3.1 Data Retrieval and Storage

Data retrieval and storage within the traditional relational model is based on indexing, but the following subsection shows through an example why it is not appropriate for spatial data. Another structure, called a quadtree, is introduced as a solution to this problem. An example is given using Oracle7 with the spatial option. See [4] for a complete reference on relational databases.

3.1.1 Indexing in Traditional Relational Databases

Traditional relational database approaches are inefficient for retrieving spatial data and answering queries that include spatial conditions because they do not take advantage of the ordering of data in two or more dimensions [5] ,[2]. For example, consider a segment of a database that contains information about various places of interest in a particular city. A segment, of the table might look as follows :

Id	Site	East	North
1	City Museum	15	60
2	Special Events	30	67
3	Sun Bow] Stadium	45	20
4	Civic Center	34	20

Now, consider the following point (example 1) and range (example 2) queries:

1. Retrieve any site at location (30,67).
2. Retrieve any site in the rectangular area defined by (10,10) and (35,70)

Using the traditional approach, it is reasonable to have two indices for the two spatial coordinate fields of our table. The indices would look as follows :

East	Site
15	City Museum
30	Special Events
34	Civic Center
45	Sun Bowl Stadium

North	Site
20	Civic Center
20	Sun Bowl Stadium
60	City Museum
67	Special Events Center

To answer the first query, we would probably do a binary search of the *East* index to locate records whose first coordinates have value 30. We then must, go to the original table to check if the second coordinates have value 67 and retrieve the records for which we have a match. For the second query, we would have to do a range search on [10,3.5] on the *East* index, where 10 refers to lower bound of the *East* index and 35 refers to the upper bound of the *East* index, giving us a list of pointers to the original table. For each pointer in the list, the specific record must be accessed and the *North* value checked in order to see if it falls in the range [10,70] for the *North* index. If it falls in that range, we retrieve the record.

The problem with this approach is that only one of the indices is used in these retrievals. This kind of indexing would be very inefficient for a large database consisting of several gigabytes of data, which is usually the case for a GIS. An indexing scheme is needed that takes advantage of the ordering in two or more dimensions.

3.1.2 Quadtrees in Spatial Databases

A structure that can be employed to overcome the problems of indexing multidimensional data is the quadtree [5],[2]. This structure is a leveled tree where all non-leaf nodes have exactly four descendants, For example, a two-dimensional region is decomposed by recursively subdividing its regions into four equal-sized quadrants. The decomposition is applied to each quadrant until the desired degree of resolution is achieved. Quadtrees are stored in a leveled-tree data structure with the root at the top level. For each non-leaf node, its four constituent quadrants are represented by its four descendant nodes. A quadrant where no further subdivisions is required is stored as a leaf node.

Oracle7 with the spatial option uses the idea of quadtrees to handle multidimensional data within the relational data model taking advantage of the customary relational query capabilities without using the traditional relational DBMS indices. To do that, Oracle7 uses an encoding technique that maintains the

dimensional organization of data. Records that reference information that are geographically near to each other are logically stored near each other. This encoding technique makes use of a new data type called the Helical Hyperspatial Code (HHCODE)[2] that allows the encoding of multiple dimensions into a unique value that is stored in a single column of a table,

The figure below shows a two-dimensional decomposition of North America. The map is divided into four equal-sized quadrants, and then each one is divided recursively into four according to what precision we want for a specific location. If a greater level of resolution is needed, the quadrants will keep subdividing. Only those quadrants in which needed data exists will keep subdividing. The four quadrants at each subdivision are given one value from 0 to 3.

The HHCODE contains a string of values from O to 3 describing the specific object that is represented according to its position and level. By level we mean the degree of decomposition, i.e., level O is the whole non-decomposed region, level 1 is the first level of decomposition, and so on. When a subdivision occurs, the number of the top-level quadrant is appended to the number of the new quadrants created by the subdivision.

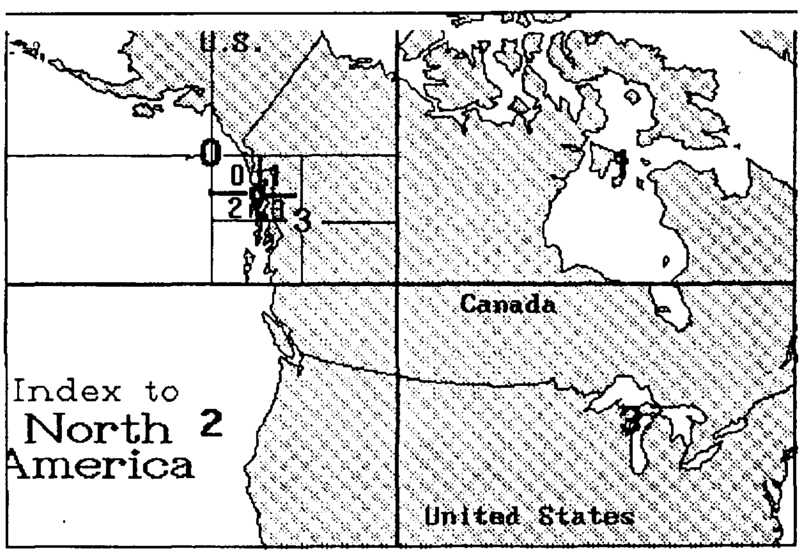


Figure 1: A decomposition of a map.

In the above figure, the data would be encoded as follows:

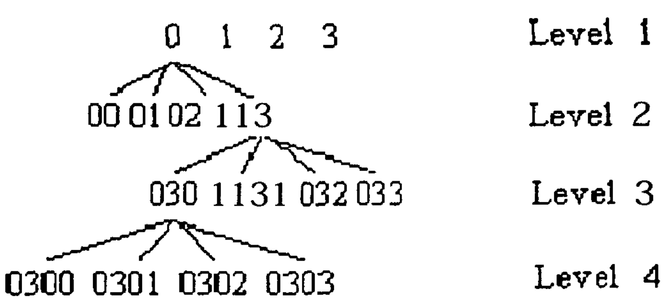


Figure 2: The quadtree representation of the decomposed map.

Note from the generated HHCODE in the example that the records containing data that are near each other geographically have common substrings. As we described in the previous section, traditional relational databases with indices on spatial data do not maintain the dimensional organization of the data. This problem is overcome by using the HHCODE, and fast retrieval of records for spatial queries is achieved.

3.2 Security

Security [5], [4],[3] is a primary concern for any database. Unauthorized access to the data should be prevented and different levels of authorized access should be allowed. For a GIS specifically, our concerns may be increased because many GIS due to their nature have distributed databases. Unauthorized access from the Internet should be prevented.

3.3 Integrity

The integrity [5],[4],[3] of data must be enforced. Data in the system should be correct and consistent with each other, e.g., data in one tile referencing an object, and data in another file referencing the same object must be consistent. Integrity over distributed data in different locations is more difficult to be enforced. Also this is true for a hybrid GIS, where data is not kept in a single database, but spatial and non- spatial data are handled separately. Precautions must be taken by the designers of the system so the integrity of the system in any case is preserved.

3.4 Data Formats

A GIS database, as we explained in section 2, must handle a wide variety of sources and types of data. The main problem here is to get the data into a format [1],[5] that. is recognizable by a particular GIS. GIS designers should be aware of the source that is **used** for capturing the data in order to avoid incompatibilities with particular data formats. Because graphic data is used extensively in a GIS, an important issue here is the space needed to store this information. Usually data in raster format occupies more space than data in vector format. Various compression techniques exist and can be employed to reduce the space required for storage of some data formats. For more on this, a good reference is [1],

4 Summary

In this paper, we gave a general background on geographical information systems and the principles on which they are based. We looked into the issues relating to the managing of data within a GIS environment, and we explained the problems with the traditional relational database approach using indexing on spatial multidimensional data. Using an example, we explained why performance with this approach is poor, We also described bow this problem can be solved using a different structure, the quadtree. An example on how this can be used was given using the Oracle7 HHCODE data structure. Also, the general principles of data security, integrity, and the handling of data formats were discussed.

Acknowledgments. This work was supported by NASA under contract NCCW-0089.

References

[1] B. Fortner., *The Data Handbook*. Santa Clara, CA: Telos, 1995.

[2] G. Gwendolyn, C. Kristian, M, Bradley, J. Rawlings., *Oracle7 Multidimension, Advances in Relational Database Technology for Spatial Data Management.* Redwood Shores, CA: Oracle Corporation, 1995.

[3] H. F. Korth and A. Silberschatz, *Database System Concepts. New York*: McGraw-Hill, 1991.

[4] J. Ullman, *Principles of Database and Knowledge-Base Systems,* Volume 1. Rockville, MD: Computer Science Press, 1988.

[5] M. F. Worboys, *GIS: A Computing Perspective.* Bristol, PA: Taylor and Francis, 1995.

[6] L. Worral, *Spatial Analysis and Spatial Policy Using Geographic Information Systems.* London: Belhaven Press, 1991.

URC97020

Off-Line Testing for Bridge Faults in CMOS Domino Logic Circuits

K. Bennett [1], P.K. Lala and F, Busaba
Department of Electrical Engineering
North Carolina A&T State University
Greensboro, NC 27411

520/61

Abstract

Bridge faults, especially in CMOS circuits, have unique characteristics which make them difficult to detect during testing. This paper presents a technique for detecting bridge faults which have an effect on the output of CMOS Domino logic circuits. The faults are modeled at the transistor level and this technique is based on analyzing the off-set of the function during off-line testing.

1. Introduction

Bridge faults are created during the design layout or manufacturing process of an integrated circuit (IC). It can result from two or more conducting paths placed too close together, the addition of extra conducting material, or insufficient insulating material. In order to minimize the number of bridge faults present in the IC, testing methods must be able to accurately detect the faults which have an effect on the normal operation of the circuit. Previous work in the area of bridge fault detection focused on static CMOS circuits.

Chess and Larrabee [1] have presented a method for generating test patterns for gate level bridge faults in static CMOS ICs. It focuses on the connection of two gate outputs which is modeled using a Fault Block and Primitive Bridge Function (PBF). The PBF represents the logic function or the behavior of the bridged components and is generated by determining whether the stimulation of the bridge fault occurs from the wire closest to the inputs of the bridged path or closest to the output.

Di and Jess [2] have presented a technique for modeling transistor level bridge faults in static CMOS circuits by evaluating the electrical behavior of the circuit and converting this behavior into logic boolean expressions called Faulty Boolean Expressions (FBE). This allows for the use of existing techniques for logic problems to determine test patterns. Gate-to-Drain type bridges are not able to be modeled.

Ferguson [3] discusses approaches for designing the physical layout of the static CMOS circuit in such a way to improve its testability of bridge faults. The three approaches are to design the circuit which reduces the number of faults, make the difficult-to-detect faults easier to detect by adding control and observation points, and make the difficult-to-detect faults unlikely to occur by considering the gate placement, circuit routing, and logic selection.

Chess, Roth, and Larrabee [4] have evaluated and compared various models used to represent bridge faults existing only between gate inputs and outputs. The different models assume that either the bridges cause wired "AND" or "OR' behavior, that the circuit value at the fault node is

[1] Currently at Motorola SPS Division -911 Battle Bend Blvd , Austin Texas

represented by a boolean function, that the analog behavior created by the fault extends beyond the fault node, or that full analog simulation is performed.

In this paper, we propose a method for detecting transistor level bridge faults in CMOS Domino Logic circuits. Only the bridge faults which are realistic at the layout level are considered. Domino logic is a type of Dynamic logic in which the on-set is used in realizing the n-logic block. Figure 1 shows a general diagram of a domino logic circuit. An invertor is connected at the output to make it low during the precharge phase. The output node *f* is precharged to "O" when the clock is low. During, the evaluation, phase i .e. when clock is high, if the input pattern closes the path between ground and output node *f*, the output is pulled to " 1" otherwise it remains at "O".

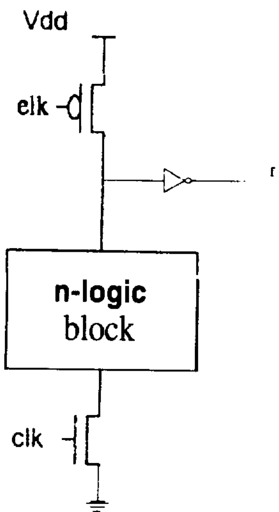


Figure 1. Domino logic circuit

Figure 2 shows a cascaded Domino logic circuit in which the output of one stage feeds the n-logic network of a subsequent stage, Only one clock is necessary for the precharge and the evaluation phase. The number of stages in the cascade depends on whether the sequence can evaluate within the evaluation cycle.

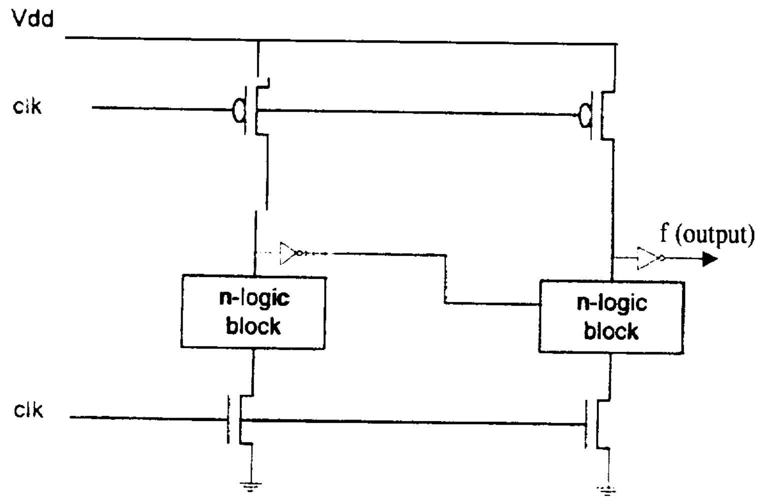


Figure 2. Cascaded domino logic circuit

The majority of the bridge faults in domino CMOS circuits occurs between the Drain/Source of one transistor and-the Drain/Source of another transistor, the Drain/Source of one transistor and the Output of the circuit, and the Drain of one transistor and Ground.

We do not consider the bridging between the Gate and Drain of the same transistor and between the gates of two **parallel** transistors .

2. Test **Pattern Generation**

As mentioned previously, we consider only bridge faults which are realistic at the layout level. The detection of these faults is based on the following lemmas.

Lemma 1 :

The output of a circuit containing bridge faults will generate more "1 's" than the fault-free circuit.

Lemma 2:

The bridge faults in a circuit can be detected by applying only the original off-set of the circuit.

To illustrate the validity of the above lemmas, we consider two examples. Example 1 considers a single stage/single output domino logic circuit. Example 2 considers a cascaded domino logic circuit.

Example 1: Figure 3 shows the implementation of the function $f = (B+ \bar{A})C + AB$. The dashed lines indicate bridge faults obtained by connecting the signal lines with metal.

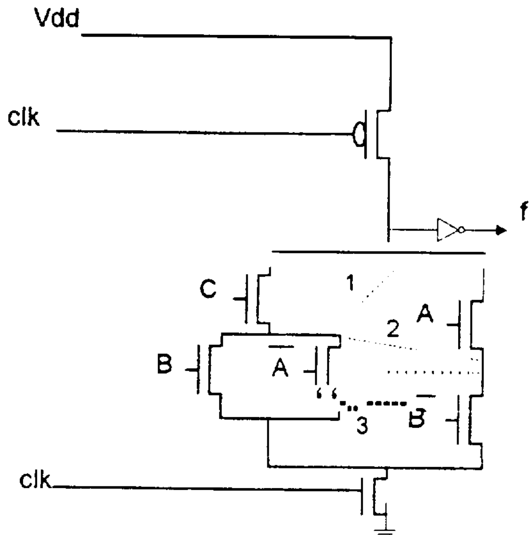


Figure 3. f= (B+ $\bar{A}$)C + A B

Table 1 shows the truth table of the fault-free and faulty functions. "A", "3", and "C" are the input variables and "*f*" represents the fault-free output. The outputs "fl_1", "fl_2", and "fl_3" represent the outputs corresponding to bridge faults #1, #2, and #3 respectively.

A	B	C	f	f_1	f_2	f_3
0	0	0	0	1	0	0
0	0	1	1	1	1	1
0	1	0	0	1	0	0
0	1	1	1	1	1	1
1	0	0	1	1	1	1
1	0	1	1	1	1	1
1	1	0	0	1	1	1
1	1	1	1	1	1	1

Table 1

As indicated by Lemma 1, the input combinations for the faulty circuits produce more "1s" than the fault-free circuit. In order to detect a bridge fault, only the input combinations which produce a "0" for the fault-free circuit need to be applied as test patterns as discussed in Lemma 2.

Table 2 shows the input combinations that detect each bridge fault.

Bridge Fault Number	Input Combination
#1	A=0 B=0 C=0 A=0 B=1 C=0 A=1 B=1 C=0
#2	A=1 B=1 C=0
#3	A=1 B=1 C=0

Table 2

Example 2: Figure 4 shows the implementation of the function f = (A+ $\bar{B}$C)D using cascaded domino logic. A bridge fault in each stage of a cascaded circuit must be tested separately. In order

to propagate the effect of the bridge fault to the final output, the inputs of the subsequent n-logic networks are set such that a "O" is produced at the outputs only if the output from the previous stage is a "O", otherwise the output is a "1". Hence, if the output of the stage under test produces a "1", this value will be propagated through the subsequent stages to the final output.

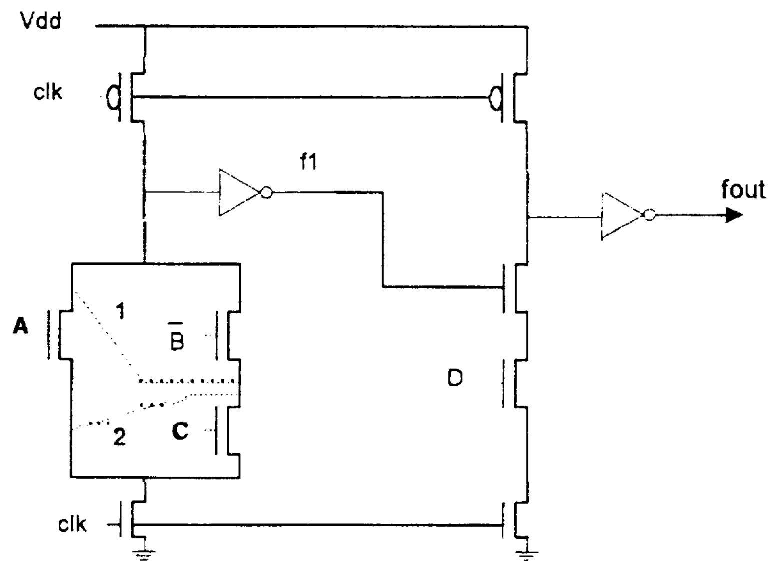


Figure 4. Cascaded domino circuit

In this example, we chose to test the first stage for the presence of a bridge fault. The value of " 1" is assigned to the input variable "D" in order to allow for the propagation of the effects produced by bridge faults # 1 and #2. Since the first stage is being tested, only its function expression needs to be evaluated (f= A+ $\bar{B}$C). Table 3 shows the input combinations that detect each bridge fault in the first stage.

Bridge Fault Number	Input Combination
#1	A=0 B=1 C=1
#2	A=0 B=0 C=0

Table 3

In order to determine the test patterns for the overall circuit, input variable "D" must be added to the input combinations shown in Table 8. Hence, the input combinations which detect bridge faults # 1 and #2 are.

A=0 B=1 C=1 D=1 and **A=0 B=0 C=0 D=1**

respectively.

5. Conclusion

We have presented-a method for detecting bridge faults in CMOS Domino logic circuits. This method is based on two lemmas discussed in section 3. This method is efficient in detecting all bridge faults except Gate-to-Drain and parallel Polysilicon bridge faults.

This technique also. applies to each stage of a cascaded circuit. The number of test patterns will be reduced corresponding to the reduction in the number of input variables used to

represent the function of each stage. In Example 3, the number of test patterns was reduced from 11 to 3 because the overall function was separated into two stages. The function expression of the stage under test consisted of only 3 input variables instead of 4.

References

[1] B. Chess and T. Larrabee. Generating Test Patterns for Bridge Faults in CMOS ICs. *Proc. The European Design & Test Conference. EDAC. The European Conference on Design Automation. ETC. Eropean Test Conference. EUROASIC, The European Event in ASIC Design,* pages 165-170. IEEE, 1994.

[2] C. Di and J.A.G. Jess. On CMOS Bridge Fault Modeling and Test Pattern Evaluation. *Digest of Papers. Eleventh Annual 1993 IEEE VLSI Test Symposium,* pages 116-118. IEEE, 1993.

[3] F, Ferguson. Physical Design for Testability for Bridges in CMOS Circuits. *Digest of Papers, Eleventh Annual 1993 IEEE VLSI Test symposium,* pages *290-295.* IEEE, 1993.

[4] B. Chess, C. Roth, and T. Larrabee. On Evaluating Competing Bridge Fault Models for CMOS ICs. *Proc. 12th IEEE VLSI Test Symposium,* pages 446-451. IEEE, 1994.

Acknowledgment: This work was supported partially by National Aeronautics and Space Administration and the Office of Naval Research under contracts ACE-48146 and ONR-N000 14-95-1-0947 respectively.

URC97021

SYSTEMS ENGINEERING CONSORTIUM FOR JPL SATELLITE URANIA

Carlos Betancourt, Juan Guillen, Irene Richardson and Dr. Bryan Usevitch
University of Texas at El Paso,
Department of Electrical Engineering

Abstract

Three students from the University of Texas at El Paso participated in the Minority University Systems Engineering (MUSE) project which was sponsored by Jet Propulsion Laboratories (JPL). The objective of the project was to introduce the concepts of systems engineering to the students by having them design the communication system for the low-orbit Satellite Urania.

1. Introduction

Systems engineering encompasses a wide range of interdisciplinary skills and an overall knowledge of system development and design in order to orchestrate the merging of all integral subdivisions of an entire project. The purpose of this endeavor was to introduce the authors of this paper to a systems engineering environment by having them design the communication system of a low-orbit satellite. This paper discusses the developmental process by analyzing the parameters which influenced the design and by formulating a composite sketch which illustrates the functions of the final design.

The spacecraft system designers designated three communication subsystems which include: a transmitter to downlink scientific instrument data, a receiver to accept command and control from the ground station, and a beacon to indicate to the ground station when the satellite is visible. Several parameters influencing the design include: the design requirements, shown on Table 1 below, which were based on the specific features of the satellite, the bandwidth of the beacon, which is calculated using the Doppler shift, and the link budget analysis, which establishes the potential system performance. The primary design of this paper focused on the transmitter and beacon subsystems.

Table 1: Subsystem Design Requirements

Subsystem	input Power	Frequency	Data Rate
Transmitter	1 watt	2.25 GHz	400 kbps
Receiver	1 watt	2.1 GHz	2 kbps
Beacon	1 watt	2.2 GHz	N/A

The basic design of the communication system is illustrated in the overall block diagram shown in Figure 1, and was initially proposed to meet the specific satellite requirements and design specifications. The diagram is comprised of a transmitter, a receiver and a beacon. The transmitter basically consists of a modulator, an RF amplifier, a bandpass filter and a patch antenna. The receiver includes a demodulator, an RF amplifier, a bandpass filter which is coupled to a four-monopole antenna. The beacon, which is connected to the four-monopole antenna is comprised of an oscillator and a bandpass filter.

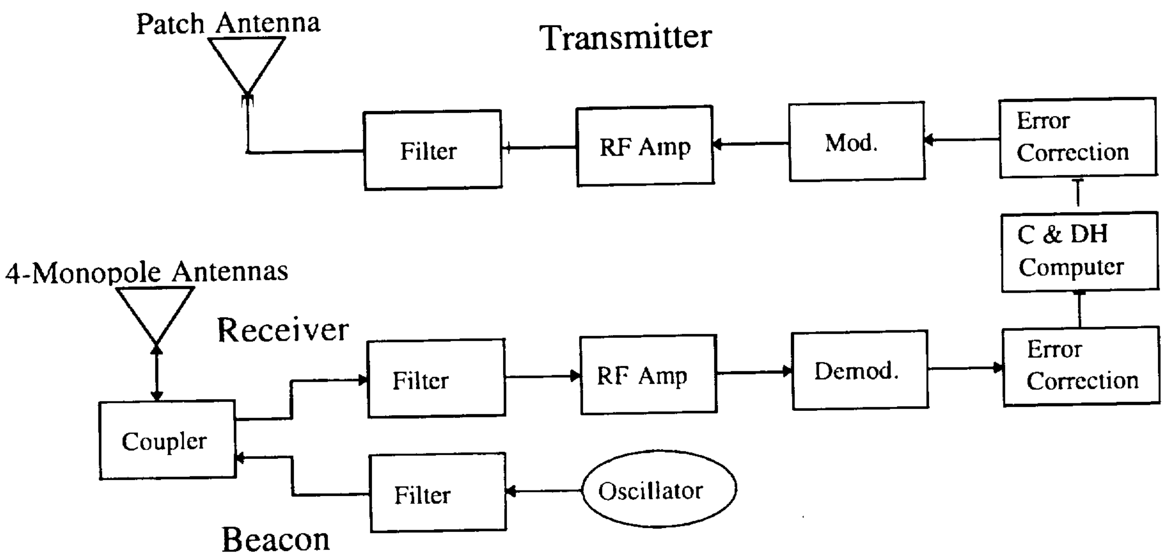


Figure 1. Overall Block Diagram

In order to properly design the subsystems, several factors need to be taken into consideration before any designing begins. The Doppler shift must be calculated to determine the bandwidth of the beacon signal. In addition, a link budget analysis must be established to calculate the power at various points of the design which is crucial in evaluating the performance of the entire communications system.

2. Doppler Effect

The Doppler effect is a phenomenon exhibited by all types of wave motion in which the apparent change in frequency is relative to the motion of the source and the observer. Since the satellite orbit was predetermined to be circular, calculations to determine the Doppler shift were straightforward using analytic geometry. Satellite parameters were needed in order to draw up a basic analysis of the geometry involving the earth and the satellite.

Ten degrees above the horizon was chosen as the angle at which the satellite first becomes visible to the ground station. Calculations for Doppler shift were made at this point since this is where Doppler shift is maximum. All calculations were done using the

Law of Sines to include distances and angles required for the Doppler Shift and Bandwidth as the satellite orbits the Earth.

The altitude of the satellite (1100 km) and the horizontal distance from the center of the earth at 38° latitude (5026 km) were critical in determining the velocity of the satellite and the earth, respectively. The angle at which the satellite transmits from its orbit above the horizon was calculated as being approximately 32.87°. Having observed these results from the analysis in the geometry, it was determined that since the velocity of the earth at that latitude was negligible compared with the velocity of the satellite, the Doppler Shift was calculated to be +42.15 kHz. The total Bandwidth for the Beacon, which includes margin for oscillator stability and drift, was assumed to be 92.2 kHz.

3. Link Budget Analysis

The satellite communication subsystem must allow us to transmit commands from the earth station and to download information from the satellite. In order to accomplish this, we need to have fidelity in the transmitted signals, and since fidelity highly depends on the signal power, we need to calculate the power at various points in the transmission and establish what is called a *link power budget.*

A link budget is necessary before the designer can start proposing components of the receiver and the transmitter. Without having a link budget, the designer would go blindly as far as how much power needs to bc transmitted from the satellite, what the size of the receive antenna should be, and other important parameters of the design. This is why a poor link budget results in poor system performance or in an expensive system.

The parameters of a link budget include the satellite orbit, radio frequencies, antennas, power amplifiers, transmission losses, noise temperature, and signal quality. For our design, the satellite is in a low circular orbit, the frequency is in the S-band, and the noise temperature is main} y caused by the Earth's temperature and galactic noise. Signal quality is achieved by having the proper signal-to-noise ratio at the receiver for demodulation and decoding of the transmitted signal.

The link budget for the communication subsystem was designed using Excel. The spreadsheet was designed to calculate the required transmitted power given the signal-to-noise ratio at the receiver, the phase-lock loop bandwidth, the system noise temperature, the Doppler shift, the oscillator stability bandwidth, the frequency of transmission, the losses, and the carrier margin. At the end, it was found that 9.47 dBm (8.84 mW) of transmitted power was required (see Figure 2).

Frequency	2.20 E+09	Hz				
Transmitted power			-20.53	dBW ==>	8.842/	mW
S/C ant. gain			0	dB	9.465	dBm
S/C EIRP			-20.53	dBW		
S/C losses			-3	dB		
Space loss			-168.7	dB		
E/S ant. gain (eff=55%), diameter:	1	m	24.65	dB		
E/S losses			-3	dB		
E/S antenna, G/kT	227.23326	dB/J				
Doppler bandwidth	1.00E+05	Hz				
Oscillator stability bandwidth	1.10E+05	Hz	-53.22	dBHz		
Boltzmann constant			228.6	dBJ/K		
System Noise Temp., T_{eq}	400	K	-26.02	dBK		
Carrier margin			-3	dB		
SNR_i			-24.21	dB		
Noise bandwidth, B_L	100	Hz				
SNR_L			6	dB		

Geometry:		
Altitude	1.10E+03	km
Elevation angle	10	deg
Nadir angle	57.13	deg
Earth angle	22.87	deg
Slant range	2950577.2	m

Figure 2. Beacon subsystem spreadsheet

4. Design Results

Once the Doppler shift has been calculated and the link budget analysis has been established, the final design can be determined by modifying the original block diagram to meet the design requirements. The final design shown below (Figure 3) uses the same 2.2 GHz high frequency sinewave oscillator for both the transmitter and beacon by incorporating a power splitter. The transmitter uses an additional oscillator at 50 MHz to bring the frequency of the 2.2 GHz oscillator to 2.25 GHz. The transmitter also consists of an I&Q modulator, two bandpass filters at both ends of an RF amplifier and a patch antenna. The beacon uses a 2.2 GHz sinewave oscillator a bandpass filter and is coupled to a 4-monopole antennas with the receiver. The receiver is comprised of a demodulator, an RF amplifier and a bandpass filter.

Communication System

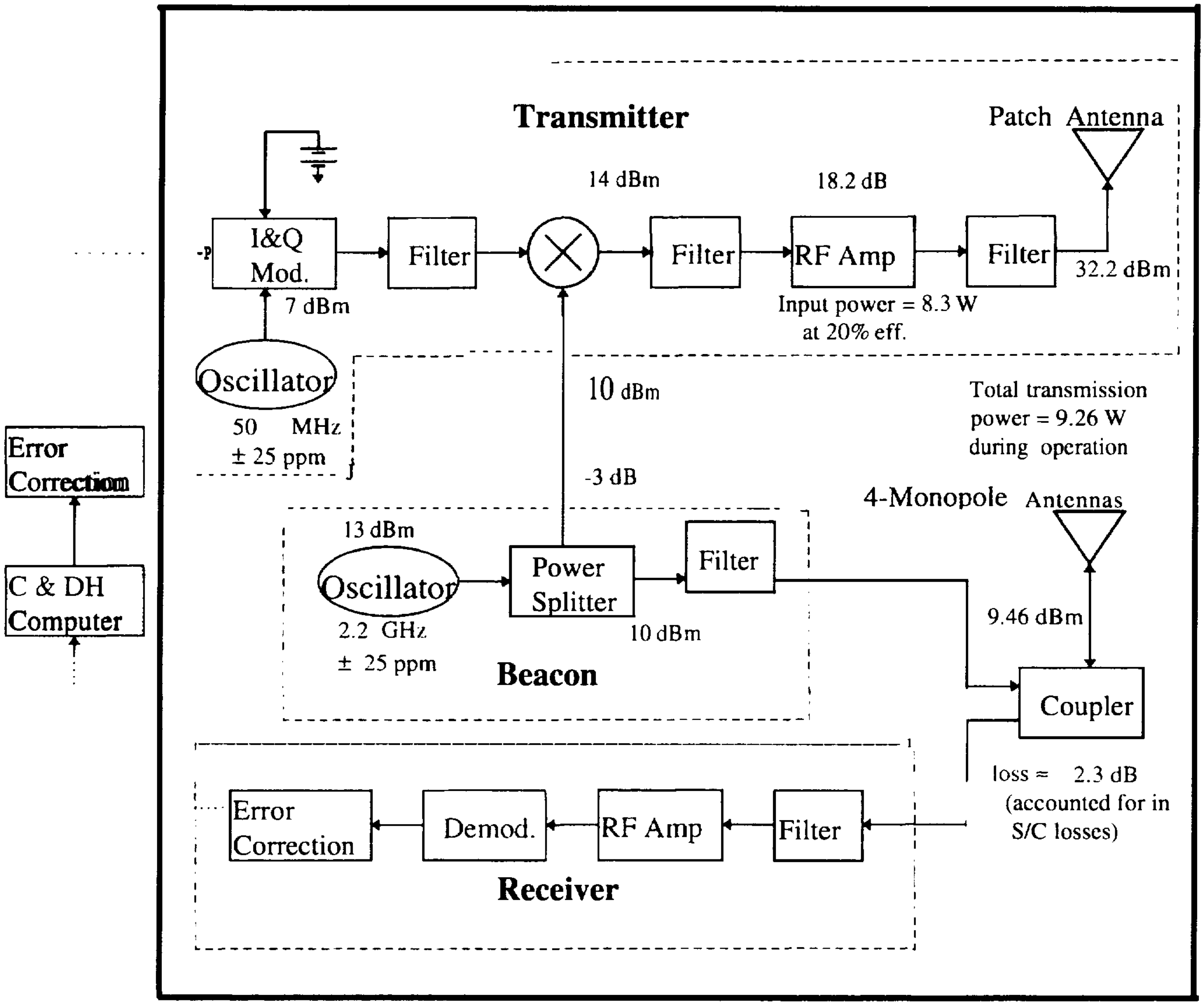


Figure 3. Final Design.

5. Acknowledgments

The students from the University of Texas at El Paso who participated in the MUSE project would like to take this opportunity to thank everyone at JPL who made this project possible. We feel that this was an invaluable experience which gave us a great insight in the field of systems engineering. We would also like to thank Dr. Bryan Usevitch for all his support and instruction.

6. References

Best, Roland E. (1984). Phase-Locked Loops. New York: McGraw-Hill.

Cheng, David K. (1993). Fundamentals of Engineering Electromagnetics. New York: Addison-Wesley.

Gagliardi, Robert M. (1991). Satellite Communications (2nd ed.). New York: Van Nostrand Reinhold.

Gardner, Floyd M. (1979). Phaselock Techniques (2nd ed.). New York: John Wiley & Sons, Inc.

Larson, Wiley J. & Wertz, James R. (1992). Space Mission Analysis and Desire (2nd ed.). Norwell, MA: Microcosm, Inc.

Miller, Franklin Jr. (1977). College Physics (4th ed.). New York: Harcourt Brace Jovanovich, Inc.

522/63

Demonstration of the Low-Cost Virtual Collaborative Environment (VCE)

David Bowers, Leticia Montes, Angel Ramos, Brendan Joyce, Dr. Ron Lumia
NASA Center for Autonomous Control Engineering (ACE)
The University of New Mexico

Abstract

This paper demonstrates the feasibility of a low-cost approach of remotely controlling equipment. Our demonstration system consists of a PC, the PUMA 560 robot with Barrett hand, and commercially available controller and teleconferencing software. The system provides a graphical user interface which allows a user to program equipment tasks and preview motions i.e., simulate the results. **Once** satisfied that the actions are both safe and accomplish the task, the remote user sends the data over the Internet to the local site for execution on the real equipment. A video link provides visual feedback to the remote sight. This technology lends itself readily to NASA's upcoming Mars expeditions by providing remote simulation and control of equipment.

Keywords: virtual collaborative environment, control, robotics, remote access, distance learning, NASA Mars expedition.

Introduction & Project Overview

The aim of this project is to develop and demonstrate the applicability of remotely controlling equipment in a cost effective manner. A remote user can program and simulate t he actions the equipment is to perform. When satisfied with their program the data are sent to the local sight for actual execution. Teleconferencing software and a camera at the local sight provide real-time viewing of the workspace to the remote user. This is demonstrated through the use of the PC, the PUMA 560 robot[1] with Barrett hand, and commercially available control ler[2] and teleconferencing [3] software.

For our application we have remotely programmed the PUMA 560 with Barrett hand to go to an object, pick the object up and move it. In this paper we will discuss the components of the low-cost VCE system, use of the system, and some applications.

Components of the Low-Cost VCE

The design of the low-cost VCE uses a PC platform. The software is broken into two components: the equipment controller software and the real-time viewing (teleconferencing) software. The equipment being controlled is the PUMA 560 robot with Barrett hand.

In this project we use the Cimetrix Open Architecture Controller. This consists of an amplifier/control module and Cimetrix controller software. The software is broken into three different components, one for program development, one for simulation and one for execution.

There were several elements to consider in determining what the data transmission and communication scheme would be. We want to provide access to University labs (the "local" site) and their equipment (robots), through an interactive environment controlled by LAN/WAN accessible computers, from a remote site computer. The local site must receive file transfers and transmit live video data. The remote site must transmit files and receive live video. A previous phase of this project[4] successfully used a remote site composed of a terminal with direct WAN access to the local site for the file transfers, and a stand alone PC with a direct phone link

(POTS) to another stand alone PC at the local site to provide a channel for video. The current phase has some unique requirements creating the need for a new solution. The first constraint is that the remote platform is a laptop running Windows NT 3.51 (necessitated by the latest Cimetrix software). Although there are several video conferencing systems available, it was difficult to find hardware and software that could run on Windows NT. Next, the only available communication channel is POTS, New Mexico does not have Integrated Services Digital Network (ISDN) capabilities. Lastly, the local LAN we are accessing does not have dial-up access.

This combination of constraints on the remote and local systems proved to be a major obstacle to finding an off-the-shelf solution. A more universal solution consists of some type of WWW accessible web site, linked to the local site, through which a remote site running a web browser would access the video feed. Due to the complex, custom programming inherent in such a configuration, we opted for a more task specific solution described below.

The teleconferencing software chosen, *Enhanced CU-SeeMe, is* made by White Pine. This software provides person to person, group conferencing, and large audience broadcasting with video, audio, chat window and white board communications over the Internet or any TCP/IP network. The camera used is called a Connectix QuickCam, which transmits video in black and white. This particular hardware was chosen because it can run on Windows NT and is compatible with *Enhanced CU-SeeMe.* The teleconferencing software and the Connectix QuickCam were relatively low in price and met the needs of the VCE system.

The resultant configuration required the addition of a LAN linked PC to the local site. The local PC, running Windows NT 3.51, is equipped with a modem (for dial in access), and a gray scale camera (QuickCam), by Connectix for the video feed. The remote site would dial into the PC, then launch a viewer application (provided by Connectix) to begin seeing the live video.

One concern is secure and safe transport of data. VCE would incorporate the use of a firewall, which requires a password and separate program for admittance. The program and password would then be electronically mailed to the remote user. Any other party trying to interrupt the transport of data will require these components. Use of VCE also requires the remote user to log into the local users system and transfer data. Another way of accomplishing this is to mail the program electronically, so that the interchanging of passwords does not occur. It must be understood by both remote and local user that robotic simulations are written in Cumulation. The *Enhanced CU-SeeMe* also incorporates security into the software, such as password usage, caller identification, and other conference and inbound call security.

The equipment being controlled is the PUMA 560 robot. This is a robotic arm containing six degrees of freedom, "Each link of the robotic arm is connected to another link at a joint, and through each joint passes one or more axes around which the links of the arm rotate. At the end of the arm is a Barrett hand used for gripping objects. The links are driven by a permanent magnet DC servomotors driving through its associated gear train. The position and velocity of each joint are needed to control the PUMA 560, and are programmed using CIMBuilder. The CIMControl software is used for real-time control of the PUMA 560. CIMControl outputs control signals to a servo board, which connects various power amplifiers to control robot motion.

Task Development

Task development is implemented via the Cimetrix Open Architecture Controller using CIMBuilder. The application is then run in Cumulation, a graphic simulation package, to preview the application. Once everything is operating correctly CIMControl is used to control the actual mechanism. Our task development goal was to use the system to remotely program the

PUMA 560 robot to pickup an object and move it following the sequence of programming, simulation, and finally execution. This proceeded as follows:

1) Build a model of the robot and the work cell

The first step in programming equipment is to make a model of the equipment and the work space in which it will operate. This is accomplished by setting nodes, geometric dimensions, and degrees of motion. The PUMA 560 contains 6 degrees of freedom and is fitted with a Barrett hand for gripping. A custom **work** cell that includes a tool table and a peg was built with CIMBuilder for use with Cumulation.

2) Program the robot

With the work cell and equipment defined the user programs the equipment to do a desired set of tasks within the work cell. This maybe accomplished in a variety of ways. One way is using CIMBuilder, a general purpose application development tool consisting of a Graphical 'User Interface (GUI) builder and a program builder allowing the user to attach a GUI to the program. Another way is to write a C program using the applications interface (API) functions from the Cimetrix Open Development Environment (CODE) Library. Here again the user programs the equipment to do a desired set of tasks. In both cases the program is built using CODE API libraries. For example to cause the robot to move to a peg, pick it up, put it down and return to its original position the user would go through a sequence similar to this:

MoveNearNode - user specifies target, tool, and distance
MoveToNode - user specifies which node or object
MoveSingleAxis - close the gripper on the peg, specify axis and value
AttachNode - attach peg to the robot gripper
MoveNearNode - move back to starting position
MoveToNode - put the peg somewhere else
MoveSingleAxis - open the gripper
AttachNode - put the peg down
MoveNearNode - return the robot to some known position

Each of these functions is defined in the CODE API libraries. All these functions may be combined **into** a procedure if desired. This procedure can then be run as a stand alone program. This procedure performs is a basic pick and place operation.

3) Simulate the program

Once either of the two programming approaches from step 2) has been done, the movement is simulated using Cumulation. This gives a simulation of the equipment's movement in the work space, allowing the user to test the program without turning on the robot. This type of simulation is quick and safe for the user. Cimetrix displays a window with a complete 3D-model of the robot and work cell. The motion and picture on the screen simulate the real robot. After the simulation is done, the user will decide on modifications or improvements if required. If no modifications are required the complete program is ready to be sent to the local site for test on the real robot.

4) Execute the program

When satisfied with the program and the results, CIMControl is **used** to move the actual equipment. Figures 1 and 2 show the task development cycle. The same program can be run with a completely different robot so long as the new robot can reach the designated object, i.e. the Peg. The simulation and control programs, Cumulation and CIMControl, are interchangeable. The application can not determine whether it is connected to the simulation or the real time robot controller,

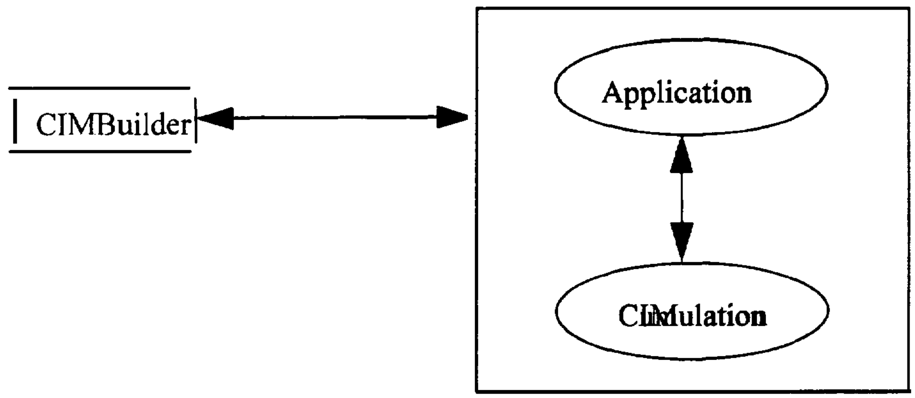


Figure 1. Using CIMBuilder to develop an application

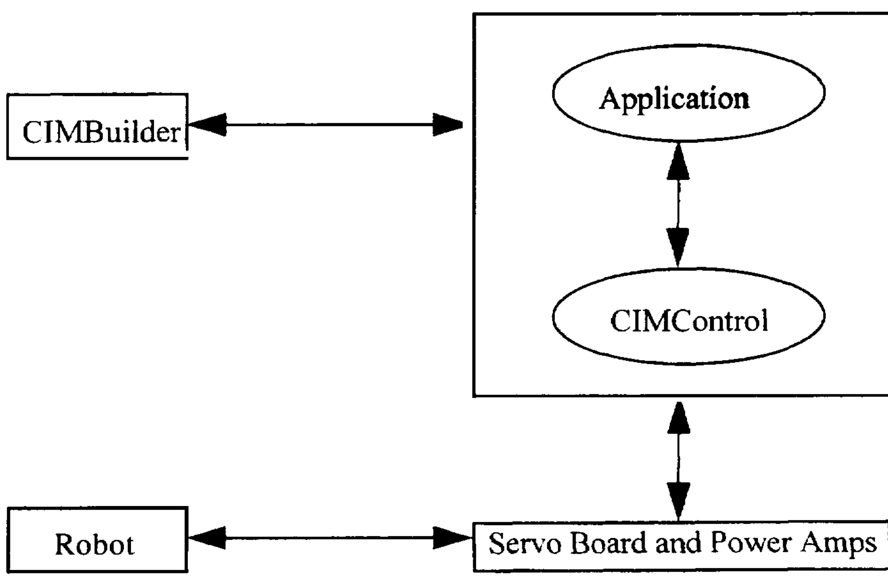


Figure 2. Real-time control of the robot

System Integration

Figure 3 shows how the pieces of the system work together. The remote site, typically using an inexpensive PC as its platform contains Cimetrix Open Architecture Controller software and White Pine teleconferencing software. The task is developed at the remote site and then sent to the local site for preview and execution. The local site contains the equipment (i.e., robot), Cimetrix Open Architecture Controller software running on any platform, and the White

Pine teleconferencing software. A camera views the work cell, sending the video to the remote site.

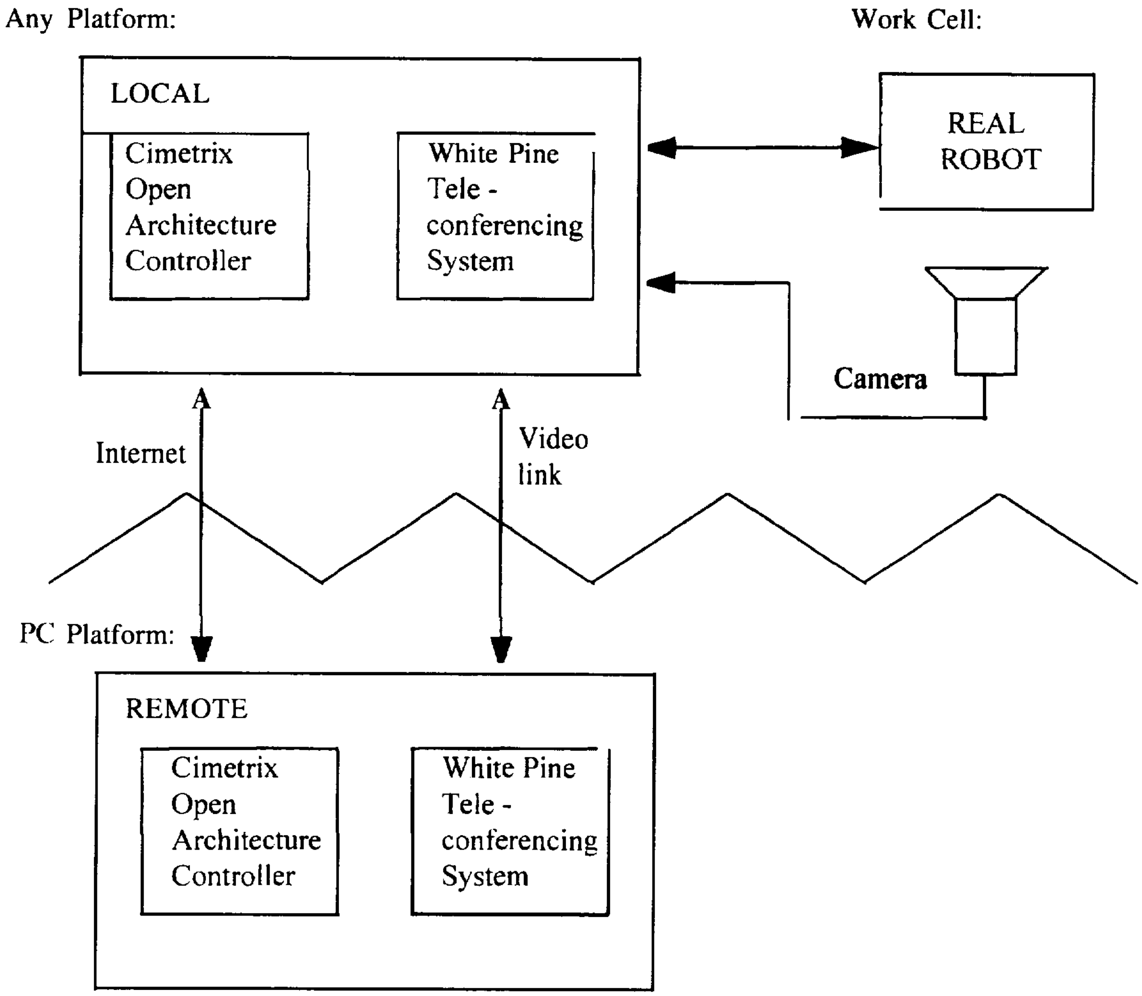


Figure 3. Low-Cost VCE System

Using the **System**

Once a work cell and equipment model have been developed using CIMBuilder, the actions are simulated, programmed and run. The data are then sent to the local site for program preview and final testing. When the local site is satisfied with the simulation the program is run. A video link provides feedback to the remote site.

Applications

Remote control of equipment has many applications. Such as:

- remote handling of contaminated waste.
- teaching factory workers how to use new equipment and simulate assembly plant actions.
- remote learning and use of equipment by rural communities, universities and public schools.
- remote control of equipment in harsh environments, e.g., space.

A space application of particular importance is the NASA Mars surveyor program.

NASA plans to send a series of spacecraft to Mars approximately every twenty-six months through the year 2005, beginning with the initial launch of two robotic explorers in November 1996. These robotic explorers are designed to study the composition of Mars, and to search for possible evidence of water. In future years, NASA envisions networks of more than a dozen of these explorers spread out over the entire planet. The Mars Surveyor Program will provide information NASA needs for the possibility of human flights to Mars, as well as create a blueprint for other planetary explorations. This program is expected to evolve as new technologies, like VCE, emerge and participate in these initiatives. VCE maybe accessed to test and preview motions of these robotic explorers to help assure capability prior to launch.

Suppose NASA scientists have received photographs of Mars, and would like to deploy a robotic explorer to search for carbonates, evaporates, or other minerals formed in the presence of water. By using CIMBuilder, the robot's task can be easily and rapidly developed using CODE. CIMBuilder is able to model the environment and terrain provided by the photographs, as well as the robots interaction with the environment, such as taking mineral samples or drilling into polar ice caps on Mars. This application would then be run in Cumulation, to graphically preview the motions of the robot. Meanwhile, NASA or a subcontractor could create a physical model of the environment. Once the remote and local users are satisfied with the simulation, the Cumulation program can be replaced with a real time control system, CIMControl, to control the actual mechanism. The video conferencing provides feedback to the remote user. Using VCE, other companies and universities collaborating in the Mars Surveyor Program may test their mechanisms, which NASA can incorporate into its mission.

The preliminary testing and motion monitoring may also optimize robot design. For instance, suppose that a design for a robot is not yet complete, or designers are not sure what the best design is for the robot task(s). VCE can test various virtual robots in different virtual scenarios, (or the same robot in different scenarios), allowing one to choose the best mechanism for its application. It can also help determine which robots currently in use would be best suited for certain applications, saving time and money.

'Future

From this base we are able to modify and add features. This includes the use of sensors, e.g., vision as part of the robot control system, an enhancement which is currently under development. This also includes the development of new scenarios and programs for possible use in the Mars Surveyor program and the development of new low level control algorithms, for example, a fuzzy logic based controller.

Conclusions

In this paper we described the feasibility of a low-cost VCE system. This was accomplished through the use of Cimetrix' integrated environment which allows a user to build a virtual environment and equipment, simulate the use of the equipment in the environment, and run the actual equipment. Cost goals were achieved by using a PC platform and inexpensive commercially available teleconferencing software.

References

[1] Unimate Puma Mark 11 Robot 500 Series - Equipment and Programming Manual, Unimation INC., Shelter Rock Lane, Danbury, Connecticut 06810

[2] Product literature, Cimetrix, Inc., 2222 South 950 East, Provo, Utah 84106,801-344-7000.

[3] Enhanced CU-Seeme Product Information, White Pine Software, INC., 1485 Saratoga Avenue, San Jose, California 95129-4934,408.446.1919

[4] Lumia, R., Todacheeney, M., Quintana, S., "Low-Cost Virtual Collaborative Environment Through Open System Technology," Proceedings of the World Automation Conference, Montpellier, France, May 1996

Acknowledgment

This work was supported in part by NASA under contract # NCCW-0087.

Page intentionally left blank

URC97023

Analysis of Vegetation and Atmospheric Correction Indices for Landsat Images

Tasha R. Bush and M. Desai
Division of Engineering

S23/43

University of Texas at San Antonio
6900 N. Loop 1604 West

tbush@paris.eng.utsa.edu

ABSTRACT

Vegetation and Atmospheric Indices are mathematical combinations of remote sensing bands which are useful in distinguishing the various values of the spectral reflectance. In this paper we study how the applications of various atmospherically corrected indices and vegetation indices can aide in retrieving the amount of surface reflectance from a remotely sensed image. Specifically, this paper studies and compares three vegetation indices and one atmospherically resistant index. These indices include the Normalized Difference Vegetation Index (NDVI), the Soil Adjusted Vegetation Index (SAVI), the Green Vegetation Index (GVI), and the Atmospherically Resistant Vegetation Index (ARVI), respectively. The algorithms attempt to estimate the optical characteristics of Thematic Mapper (TM) imagery. It will be shown that the NDVI algorithm followed by the ARVI correcting algorithm provided significant improvements in the tonal qualities of the retrieved images. The results are presented on 1987 TM images over the Kennedy Space Center (KSC) and are compared with a set of United States Geological Survey (U. S.G.S) maps.

L INTRODUCTION

The calculation of vegetation densities from remotely sensed imagery has been used for developing and validating various studies regarding land cover dynamics such as global carbon modeling, biogeochemical cycling, hydrological modeling, and classification response modeling [1]. However, the remote imagery collected by satellites are often contaminated by the effects of absorption and scattering of radiation from the earth's surface by surrounding atmospheric particles. This scattering effectively reduces the amount of radiation in a given scene, and therefore reduces the amount of surface reflectance from the scene. For any given material, the amount of radiation that is reflected from a substance will vary with wavelength [6]. This important property of matter allows for the possibility that different substances or classes can be identified and separated by their spectral signatures. Therefore at certain wavelengths, green vegetation may reflect more light than water or sands, however at another wavelength it absorbs more light and therefore has a reduced reflection. One method that is useful in differentiating the various values of spectral reflectance in a remote image is the application of vegetation and atmospherically resistant indices [5].

The purpose of this paper is to investigate the usefulness of three basic vegetation indices for KSC images and one atmospherically resistant index: the Normalized Vegetation Index (NDVI), the Soil Adjusted Vegetation Index (SAVI), the Green Vegetation Index (GVI), and the Atmospherically Resistant Vegetation Index (ARVI) in the mapping of various vegetation around the Kennedy Space Center. The images are of dimensions of 512 rows by 512 columns.

The resulting vegetation maps were compared with a set of United States Geological Survey (U. S.G.S) maps, in order to determine how effective each algorithm proved in determining the amount of vegetation that actually resided in the geographic structures. The

U. S.G.S. map serves as the ground truth and displays the regions of vegetation and other landcover materials.

The Normalized Difference Vegetation Index (NDVI) was found using the following equation[4]:

$$NDVI = (TM4-TM3)/(TM4+TM3),$$

where TM3 and TM4 represent the reflectance **values** of a pixel in the TM band 3 and the TM band 4 images, respectively. In the image bands, the numerical pixel value (O to 255) was taken to be the corresponding reflectance value. The NDVI algorithm was applied on a pixel by pixel basis to the TM bands in order to obtain the resulting set of TM data. When evaluating the NDVI results, a large reflectance value is represented by a high concentration of vegetation in the corresponding geographic area. Therefore, the higher the pixel intensity value, the more vegetation coverage was found to exist within that individual pixel.

The Soil Adjusted Vegetation Index (SAW) was applied to the TM bands in a similar fashion. The SAW images were based on the following equation[1]-[2]:

$$SAVI = \{(TM4 - TM3)(1+L)\} / \{TM4 + TM3 + L\},$$

where L is a correction factor which ranges from O for very high vegetation coverage to 1 for very low vegetation coverage [1]. The (1 + L) multiplicative term is used in order for the range of the vegetation index to be from 1 to -1. This is done so that SAVI reduces to NDVI when the adjustment factor goes to zero. A value of 0.5 was used for this research since it was considered typical for intermediate vegetation coverage.

The Green Vegetation Index (GVI) was also applied to the various TM 1, TM2, TM3, TM4, TM5, and TM7 bands. The Green Vegetation Index is an algebraic function combining the previously listed TM bands. The equation for evaluating the resulting GVI data set is given by[1]:

$$GVI = (-0.2848*TM1) - (0.2435*TM2) - (0.5436*TM3) + (0.7243*TM4) + (0.0840*TM5) - (O.180*TM7).$$

The Atmospherically Resistant Vegetation Index (ARVI) is one of a family of built-in atmospheric correcting indices. The form of the equation for ARVI is similar to that of NDVI with the exception that the TM band 3, the red reflectance band, in NDVI is replaced with a term combining TM 1, TM3, and γ. The ARVI images were found in a corresponding fashion based on the following equation [4]:

$$ARVI = (TM4 - rb)/(TM4 + rb),$$

with rb defined as:

$$rb = TM3 - \gamma(TM3 - TM1).$$

The value of y was varied between 0.5 and 1.0. However, only the results using γ=1.0 are presented here.

II. NDVI RESULTS

The water covered areas of the NDVI image were generally represented by low reflectance values, near black in color. The NDVI image revealed that some parts of the water contained some amount of vegetation within the pixels. Possible reasons for the lighter areas are due to shallow waters consisting of underwater vegetation or floating vegetation. Also, some of areas mapped as water by the U .S.G.S. were only inundated part of the year. The vegetated areas of the NDVI image are represented by the medium gray levels. The areas of high vegetation on the NDVI image did correspond well to the mapped vegetation of the U.S.G.S. map. The highest

intensity of pixels, near white in color. represent areas of scarce or zero vegetation. These areas of zero vegetation include sands, highways, and buildings, see Figure 1.

The SAVI image contained similar comparisons to that of the NDVI image. The SAVI corrected image again represents water coverage with low intensity pixels (near black). In some of the water areas, the SAVI image was not able to discern the neighboring land pixel intensities. However, unlike the NDVI image, the vegetated areas of the SAVI image indicated intensity levels where less vegetation occurred. These intensity levels of the SAVI image could be closely matched to that of the U.S.G.S. maps which indicate that these areas contain more soil and are less vegetated, see Figure 2.

The GVI image resulted in pixel intensity values in which some of the areas of water could not be discerned with that of neighboring land pixels. The GVI image was not useful in determining the highly vegetated areas from areas which contained medium vegetation coverage, see Figure 3.

The ARVI results with a gamma (γ) representation of 1.0 showed vegetation and water coverage areas similar to that of the NDVI image. The areas of zero vegetation exhibits higher pixel values similar to NDVI. However, errors did occur in the water areas. The water areas of the ARVI image tend to have higher intensity levels than that of the NDVI image, see Figure 4.

VI. CONCLUSIONS

In general, this project accomplished its goal of using vegetation and atmospherically resistant indices as a unit of measure for recovering and examining the reflectance values of vegetation in the geographic region of interest. NDVI appeared to distinguish vegetation, sands, and water areas fairly well. This study also determined that ARVI also proved well in determining vegetated and water areas. SAVI was useful in determining the less vegetated lands, however, some of the water areas and the neighboring land areas had similar pixel intensities and were difficult to discern between the two. The GVI image appeared not to be useful for this geographical area under study.

ACKNOWLEDGMENTS

This project was partially funded by a grant from NASA, Project #NAG10-0 155.

REFERENCES

[1] Baret, F. And Guyot, G. "Potentials and Limits of Vegetation Indices for LAI and APAR Assessment," *Remote Sensing Environment,* vol .35, pp. 161-173.

[2] Huente, A. R., H.Q. Lui, " An Error and Sensitivity Analysis of the Atmospheric and Soil Correcting Variants of the NDVI for MODIS-EOS," *IEEE Transactions on Geoscience and Remote Sensing,* vol. *32,* no. *4,* pp. 897-905, July 1994.

[3] Jensen, John R., Introductory Digital Image Processing- A Remote Sensing Perspective, 2nd edition, Prentice Hall, New Jersey, 1996.

[4] Kaufman, Y. J., Tanre, D., "Atmospherically Resistant Vegetation Index (ARVI) for EOS-MODIS," *IEEE Transactions on Geoscience and Remote Sensing, vol. 30,* no. *2,* pp. 261-270, March 1992.

[5] Lillesand, T.M. and Keifer, R.W. Remote Sensing and Image interpretation, 2nd edition, John Wiley and Sons, New York, Chichester, 1987.

[6] Sabins, Remote Sensing Principles and Interpretation, 2nd edition, W.H. Freeman and Company, New York, 1978.

Figure 1: NDVI Image

Figure 2: SAVI Image (L=0.5)

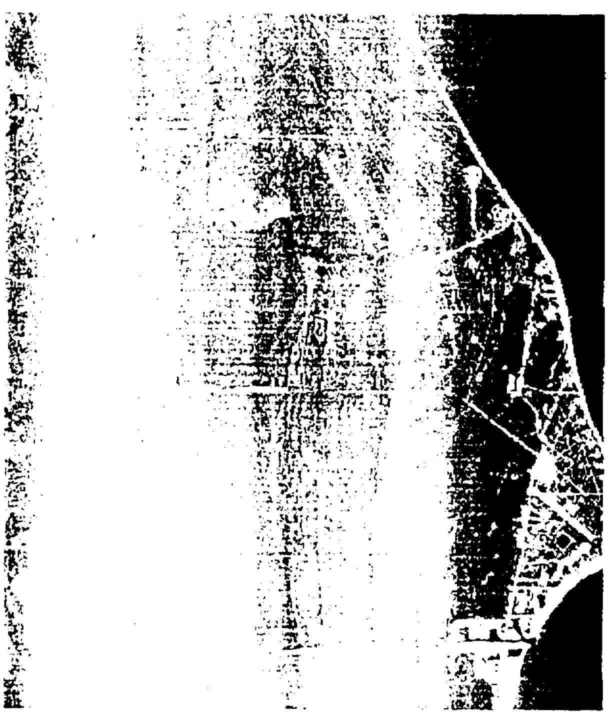


Figure 3: GVI Image

Figure 4: ARVI Image (γ=1.0)

URC97024

A Simplified Plasma Model for Reactor Design Applications

34/75

Chantay Canty, Sy-Chyi Lin and Jorge Gabitto
Chemical Engineering Department
Prairie View A&M University

ABSTRACT

Due to the increasing use in the microelectronics industry of plasma processes simpler methods to design and operate plasma etching and plasma deposition reactors are needed. Traditional plasma models require use of high performance computers to predict the influence of operating variables on plasma properties. In order to obtain a simpler design of plasma reactors a simplified model is presented. The model is based **on** *standard literature assumptions and experimental information for dilute plasmas. The main plasma properties, mean electron density, mean electron temperature and rate of plasma reactions are calculated from reactor geometry and operating variables, Good agreement was observed by comparison between calculated* results and experimental *data.*

INTRODUCTION

Plasma-assisted etching and deposition of thin films using reactive glow discharges is currently use widely in the microelectronics industry [1]. Several models to compute plasma properties has been published in literature ([2] to [5]). Most of them compute electron temperature and density as function of geometry and operating variables by solving Boltzmann equations for the different charge species.

Rogoff et al. [3] presented a model that describes the electrical characteristics of the bulk plasma region in a 13.56 -MHz parallel plate discharge in low pressure chlorine gas. They modeled the plasma as a volume-controlled region with the electron balance dominated by the single step electron impact ionization and attachment and with the electron energy distribution function in equilibrium with the local instantaneous electric field. Solutions of the Boltzmann equations for mixtures of Cl_2 and Cl which result from Cl_2 dissociation were reported. Kushner [4] studied the discharge kinetics and plasma chemistry during plasma deposition of amorphous silicon. He used a plasma model based whose electron impact rate constants were obtained from a Monte Carlo simulation for the electron distribution function in parallel plate rf discharges.

Park and Economu [5] presented a model of a chlorine discharge based on solutions of the charge species balance assuming several collision processes in the gas phase. Economu et al. [6] determined the main plasma properties using a stochastic method to simulate the different elastic, inelastic and excitation processes in a plasma. This research involved massive use of computer power. Most of the aforementioned research relies heavily on extensive use of high performance computers. These methods ,while justified from the academic point of view, are not normally use in industry. Industrial practitioners design plasma reactors mostly using experimental information ([7], [8]). Therefore, simpler methods to calculate the main plasma properties, mean electron temperature and density and rate of plasma reactions, are needed.

The goal of this research is to derive a set of formulas that will allow us to calculate ,within engineering precision, the properties of the discharge plasma from the system design and the operating parameters.

THEORETICAL MODEL

A plasma is a partially ionized gas. It consists of ion, radicals, electrons and neutral molecular species. A plasma of the type of interest to us here is usually created at low pressure, typically 0.001 to 1 torr, and consists primarily of neutral molecules. The presence of a small fraction of high energy electrons produces collisions with neutral molecular species that can dissociate into radicals. These reactive radicals can participate in rapid chemical reactions at low

temperatures. When the reactive species in the plasma, primarily the neutral radicals, combine with the atoms of a solid surface film, volatile products can be formed leading to etching. Deposition fragments can be also formed leading to deposition.

The electron distribution, the electron temperature and the specific rate of plasma reactions as a function of reactor geometry and operating variables are needed for successful reactor design.

We start by considering a plasma confined within a cylindrical reactor (see Figure 1). Under this assumption the electron diffusivity, De, is given by,

$$De= vi \Lambda^2 \quad (1),$$

where ,v_i , is the inelastic collision frequency and ,Λ, is the electron diffusion length given by (Park and Economu [5]),

$$\frac{1}{\Lambda^2}=\frac{2.405}{R^2}+\frac{\pi}{H} \quad (2),$$

here R and H are the reactor radius and the electrode separation, respectively.
The inelastic collision frequency can be calculated using (Barrington [9]),

$$v_i = N\ v_d\ \sigma_i \quad (3),$$

where N is the chemical species concentration in the gas phase, v_d is the drift velocity and σ_i is the process collision frequency.

We assume that the process collision frequencies, vi ,and ,v_m , the momentum collision frequency are constant independents of the drift velocity. This is a crude approximation, but it will allow us a considerable simplification in our derivation.

We are going to use the following equation valid for gases formed by simple molecules such as, helium and hydrogen (Barrington [9]),

$$v_d = C_1 \frac{E_{eff}}{p} \quad (4),$$

where ,C_1, is a constant ,E_{eff}, is the effective electric field and ,p, is the gas pressure.

We are going to assume that the electron in the gas phase can be represented by a Margenau distribution (Hollahan and Bell [10]). This assumption considers all the electrons to be at the same temperature, but unlike the Boltzmann distribution the electron temperature ,T_e, is different from the gas phase (T). For a Margenau distribution the electron temperature can be calculated from,

$$k\ T_e = \frac{e^2\ E_{eff}^2\ M}{6\ m^2\ v_m^2} \quad (5),$$

where k is the Boltzmann constant, e is the electron charge, m the electron mass and, M is the mass of the gas main chemical species. We use the fact that the electron diffusivity and the electron temperature are related by (Hollahan and Bell [10]),

$$D_e = \frac{k\ T_e}{m\ v_m} \quad (6),$$

in order to combine equations (5) and (6),

$$D_e = \frac{e^2\ E_{eff}^2\ M}{6\ m^3\ v_m^3} \quad (7),$$

from equation (l),

$$v_i = \frac{D}{\Lambda^2}\ \frac{e^2\ E_{eff}^2\ M}{6\ A\ m^3\ v_m^3} \quad (8),$$

or by using equation (3),

$$v_i = \frac{e^2\ E_{eff}^2\ M}{6\ \Lambda^2\ m^3\ v_d^3\ N^3\ \sigma_m^3} = N\ v_d\ \sigma_i \quad (9),$$

or combining (9) and (4),

$$\frac{e^2\ E_{eff}^2\ M}{6\ \Lambda^2\ m^3\ v_m^3} = N^4\ v_d^4\ \sigma_i = \{N\ C_1 \frac{E_{eff}}{p}\}^4\ \sigma_i \quad (lo),$$

using the ideal gas approximation,

$$N = \frac{p}{k\ T} \qquad (11),$$

we can rearranged equation (10) as,

$$\frac{e^2\ M}{6\ \Lambda^2\, m^3\, v_m{}^3\, C_1{}^4\, \sigma_i} = \frac{p^2}{k^4\ T^4}\frac{E_{eff}{}^2}{p^2} \qquad (12),$$

or,

$$\frac{E_{eff}}{p} = \{\frac{e^2 M\ \ k^4\, T^4}{6\ \Lambda^2\ \ m^3\ \ v_m{}^3\ \ C_1{}^4\, \sigma_i}\}^{(1/2)}\ \frac{1}{p\,\Lambda} \qquad (13),$$

Equation (13) allows us to calculate the effective electric field per unit pressure as a function of electron constants, gas phase temperature and mass and constants (σ_m,σ_i,C_1).

The power dissipated by unit plasma volume is given by (Hollahan and Bell [10]),

$$\frac{P}{V_r} = \frac{n_{eav}\ e\ E_{eff}{}^2}{m\ v_m} \qquad (14),$$

where n_{eav} is the average electron density, P is the power and V_r the reactor volume is given by,

$$V_r = \pi\, e^2\, H \qquad (15),$$

combining equations (3), (4) and (14) we get,

$$n_{eav} = \frac{P\ m\ C_1\ v_m}{e^2\ E_{eff}\ k\ T} \qquad (16),$$

the following relations for a cylindrical reactor reported by Hollahan and Bell [10] lead to,

$$n_{eo} = 3.64\ n_{eav} \qquad (17),$$

and

$$n_e = n_{eo}\, J_O\, \{A\ r/R\}\ \cos(B\ z/H) \qquad (18),$$

here n_{eo} is the electron density on the reactor axis, J_O is the Bessel function of order O and r and z are the radial and axial coordinates, respectively.

The reaction coefficients for plasma reactions (k_p) can be calculated using the following equation reported by Bell [10},

$$k_p = \sqrt{\frac{8}{\pi\ m}}\,(k\ T_e)^{-3/2} \int_0^{\infty} \varepsilon\ \sigma_i(\varepsilon)\ \exp(-\frac{\varepsilon}{k\ T_e}\ d\varepsilon) \qquad (19),$$

here ε is the electron energy and $\sigma_i(\varepsilon)$ is the reaction cross section,

Equation (19) can be integrated with the additional assumption that $\sigma_i(\varepsilon)$ is not a function of ε. Under this circumstances the following equation is derived,

$$k_p = \sqrt{\frac{8}{\pi\ m}}\ (k\ T_e)^{1/2}\, \sigma_i \qquad (20).$$

Equations (13), (16) and (20) allow us to calculate the main plasma properties, $\frac{E_{eff}}{P}$, electron density distribution and the reaction coefficients for plasma reactions.

RESULTS AND DISCUSSION

In order to implement the model described above the following procedure is recommended. First, calculate A from equation (2), then from p, A and data values calculate E_{eff}/p from equation (13). Equation (16) is used to calculate n_{eav}. From equations (17) and (18) the electron density distribution is computed in a reactor shown in Figure 1. The electron temperature (T_e) is calculated using equation (5). This temperature is used in equation (20) to calculate the specific reaction rate constants.

Figure 2 shows the calculated values of E_{eff}/p at different value of $p\Lambda$ for the breakdown of silane between two parallel electrodes. The calculated results are close to those reported by Bell

[10]. At low values of $p\Lambda$, E_{eff}/p is quite large due to the rapid loss of electrons under these conditions, As $p\Lambda$ increases, E_{eff}/p decreases and tends to level out. These characteristics are due to the reduced loss of electrons by diffusion.

The average electron densities at different $p\Lambda$ are shown in Figure 3. This figure illustrates the behavior of $n_{e0}/\langle P\rangle\Lambda$ vs. $p\Lambda$. Variation in the pressure, for a constant value of E_{eff}/p, will affect only the values of $p\Lambda$. When the gas pressure p increases, it causes a decrease in the electron density n_e needed to maintain a fixed value of $\langle P\rangle$. The result is very close to Bell's result [10]. The maximum error is less than 50%.

The electron density distribution at different radial position is shown in Figure 4. The electron density is seen to decrease when either r* or z* is increased. This means that the electron density value at the mid-plane of the discharge is higher than the values at the other positions. We can also see that the electron density decreases slowly along the radial direction. This fact allows us to consider a one-dimensional plasma model for describing the case of a reactor with flow.

Figure 5 shows a plot of T_e (eV) as a function of E_{eff}/p. The electron temperature is known to be dependent upon the electric field strength E_{eff} and the gas pressure p. The electron temperature is related to the mean electron energy (ε). This means that increasing the electric field strength E_{eff} or decreasing the gas pressure p, will increase electron's energy (or high temperature). Neutral molecules are easily excited when the electron temperature is high. In Figure 5, We also compare the calculated results with the literature ones [10]. It is found that the two set of data at low values of E_{eff}/p are very similar, the literature data become flat when the values of E_{eff}/p increase, but the value of the calculated data still increases.

Values of dissociation rate constant k_p computed from Eq.(19) are shown in Figure 6 as a function of E_{eff}/p. Increasing E_{eff}/p means that the electron energy is increased. The possibility of dissociation reaction between electrons and neutral molecules is increasing. Therefore, the dissociation rate constant k_p increases with a increase of E_{eff}/p. The calculated results are very similar to the literature reported results [10].

CONCLUSIONS

A simplified model to calculate plasma properties from reactor geometry and operating variables has been proposed. Dilute plasma, linear dependence of drift velocity on E_{eff}/p, v_m independent of electron velocity and the electron density function following a Margenau distribution are the main assumptions. These assumptions lead to a set of formulas that can be used to determine E_{eff}/p, T_e, and the plasma reaction specific coefficients (k_p) from the geometry of the reactor and operating variables. The calculated values were satisfactorily compare to experimental results.

REFERENCES

1). Sherman, A., "Chemical vapor deposition for microelectronics." *Noyes publications,* Park Ridge, New Yersey, USA (1987).

2). Economu, D. J.; Park, S. K. and Williams, G. D., *J. Electrochem. Sot.* , **136**, 188(1989).

3). Rogoff, G.; Kramer, J. and Piejak, R., "A Model for the Bulk Plasma in an RF Chlorine Discharge." *IEEE Trans. Plasma Sci.,* **PS-14**, *103(1986).*

4). Kushner, N. J., "A Model for the Discharge Kinetics and Plasma Chemistry during Plasma Enhanced Chemical Vapor Deposition of Amorphous Silicon." *J. Appl. Phys.*, **63**, 2532October (1988).

5). Park, S-K, and Economu, D. J., "Numerical Simulation of a Single -Wafer Isothermal Plasma Etching Reactor." *J. Electrochem. Sot.* , **137**, *2624(1990).*

6). Economu, D. J.; Wise, R. and Lymberopoulus,D., "Rarified Plasma Flow Predicted by Direct Simulation Monte Carlo Method," *Proceedings AIChE Annual Meeting, San* Francisco, CA, November (1994).

7). Middleman, S. and Hochberg,A. K., "Process Engineering Analysis in Semiconductor Device Fabrication." McGraw Hill (1993).
8). Lee, H.H., " Fundamentals of Microelectronics Processing", McGraw Hill (1990).
9). Barrington, B. E., "Gaseous Electronics and Gas Laser." Pergamon Press, Oxford (1979).
10). Hollahan, J. R. and Bell, A. T., "Techniques and Applications of Plasma Chemistry. " John Wiley & Sons, New York (1974)

Figures

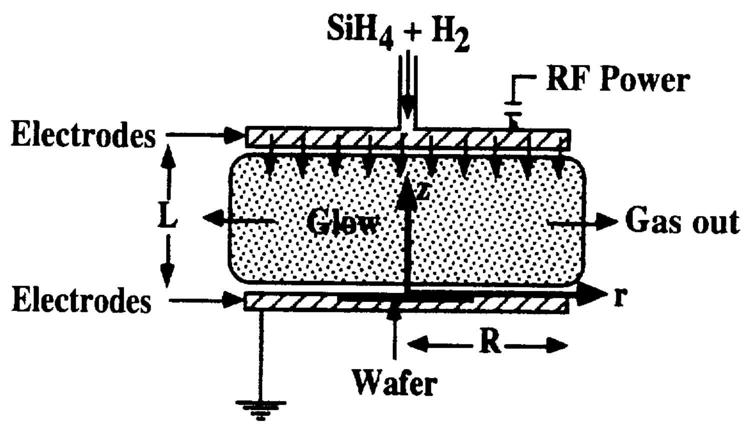


Figure 1. A sketch of the reactor of the plasma-enhanced chemical vapor deposition process.

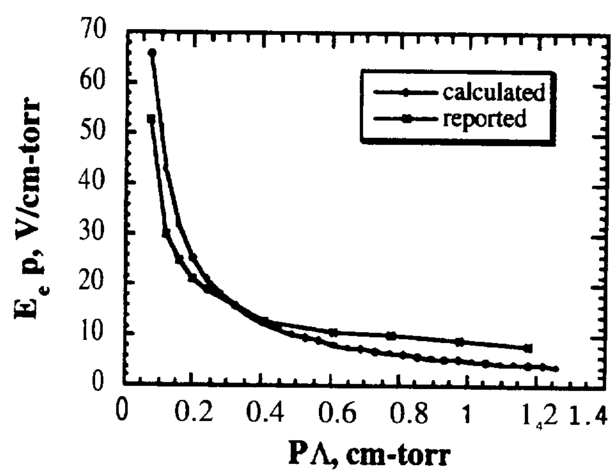


Figure 2. The value of Ee/p at different value of pΛ.

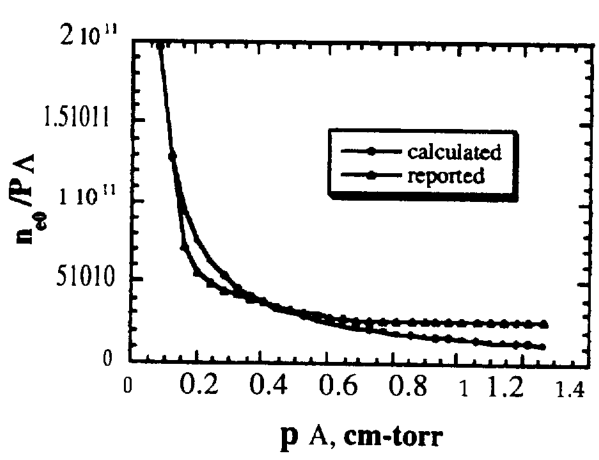


Figure 3. The average electron density calculated at different pΛ value.

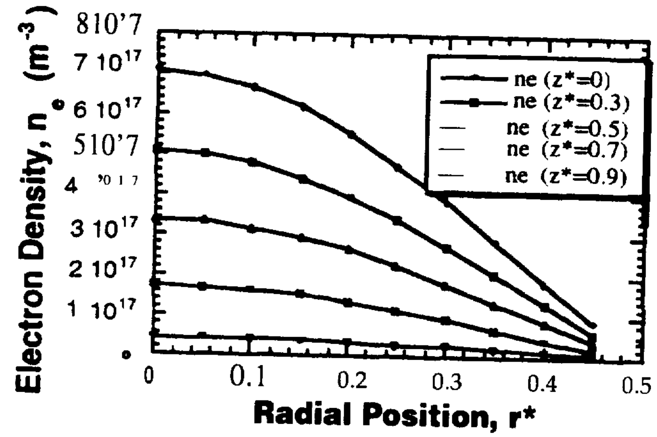


Figure 4. The distribution of electron density at different radial position.

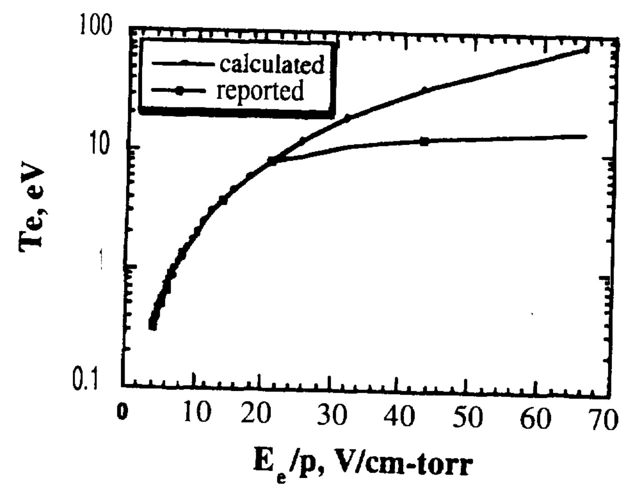


Figure 5. The electron kinetic energy at different value of Ee/p.

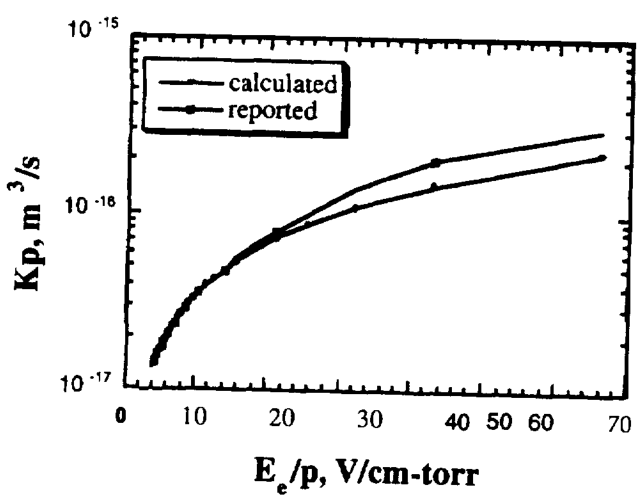


Figure 6. The dissociation rate constant of silane at different value of Ee/p.

URC97025

CREW MEMBER INTERFACE WITH SPACE STATION FURNACE FACILITY

S25/54

Martha B. Cash
Crew Systems Branch
NASA/Marshall Space Flight Center

Introduction

The Space Station Furnace Facility (SSFF) is a facility located in the International Space Station United States Laboratory (ISS US Lab) for materials research in the microgravity environment. The SSFF will accommodate basic research, commercial applications, and studies of phenomena of metals and alloys, electronic and photonic materials, and glasses and ceramics. To support this broad base of research requirements, the SSFF will operate, regulate, and support a variety of Experiment Modules (EMs). To meet station requirements concerning the microgravity level needed for experiments, station is providing an active vibration isolation system, and SSFF provides the interface.

SSFF physically consists of a Core Rack and two instrument racks (IRs) that occupy three adjacent ISS US Lab rack locations within the International Space Station (ISS). This generic SSFF configuration is shown in Figure L All SSFF racks are modified International Standard Payload Racks (ISPR). SSFF racks will have a 50% larger pass through area on the lower sides than ISPRs to accommodate the many rack to rack interconnections. The Instrument Racks are further modified with lowered floors and an additional removable panel (15" x 22") on top of the rack for access if needed. The Core Rack shall contain all centralized Core subsystems and ISS subsystem equipment. The two Instrument Racks shall contain the distributed Core subsystem equipment, ISS subsystem equipment, and the EMs. The Core System. which includes the Core Rack, the IR structures, and subsystem components located in the IRs serves as the central control and management for the IRs and the EMs. The Core System receives the resources provided by the International Space Station (ISS) and modifies, allocates, and distributes these resources to meet the operational requirements of the furnaces. The Core System is able to support a total of four EMs, and can control, support, and activate/deactivate the operations of two EMs simultaneously. The IRs can be configured to house two small EMs or one tall vertical EM, and serve as the interface between the Core and the respective EM.

The Core *Rack* and an adjacent Instrument Rack **(containing one or more furnaces) will be delivered to the ISS in one** launch. This is Integrated Configuration One (IC1). The Core Rack and IR1 will be passive *during transport* in the Mini Pressurized Logistics Module (MPLM). Any subsequent EMs to operate within IR1 are installed on-orbit. The second IR (containing one or more furnaces) is delivered to ISS on a subsequent launch which will establish Integrated Configuration Two (IC2). **Additional integrated configurations will be established with the replacement** of EMs or Instrument Racks.

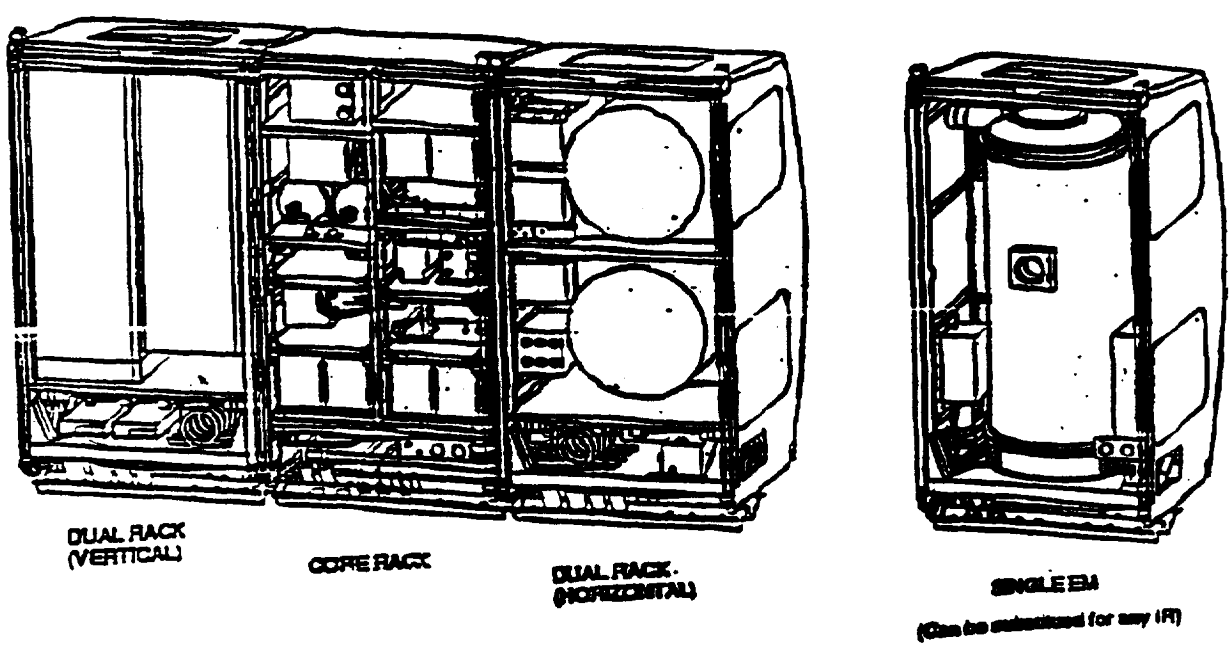


Figure 1. SSFF Generic Configuration

SSFF/Crew Interface

Initially, the physical layout, launch configuration, and basic functions of SSFF must be understood *in* order to identify crew interfaces necessary for operation. Crew time available is baselined at 90 crew hours per week for all payloads. Assuming that the Lab has 12 ISPRs, 6 ISPRs in the Attached Pressurized module, and 12 ISPRs in the Japanese Experiment Module, the result is about three crew hours per week per rack on average. This is a small number which must include installation, maintenance, and routine sample exchange. Therefore, SSFF has chosen a high degree of automation for most of its functions and considered the following critical crew interfaces in order for the facility to make efficient use of crew time Visual and reach envelopes, Operating forces, Crew and loose equipment restraints, Data displays and controls, Microgravity and the resulting Neutral Body Posture, General and task lighting, Tools, Fasteners, Connectors, Crew Safety, Workstations, Serviceability, and Labeling.

A brief description of some SSFF (based upon current configuration) related tasks which depend upon crew members are as follows: The Core Rack will be launched with all its components in place and six additional flight boxes intended for on orbit installation into the instrument racks. The Core Rack will be moved by the crew from the Logistics Module to its permanent operating location inside the Lab. After installing the Core Rack into position, the crew returns to the Logistics Module and transports Instrument Rack 1 (IR1) into the Lab and into position on the left hand side of the Core Rack. IR1 is launched with one Experiment Module (EM) already installed but requires the on orbit installation of four flight boxes. These must be removed from their bolted down launch position inside the Core Rack and placed in operational location in IR1; securing the fasteners and connectors of each &vice. The other two devices remain stowed in the Core Rack until the arrival of IR2, at which time the two remaining boxes will be installed in their operational position. The unattached harness cables must be restrained and all exposed connectors must be protected with covers, or attached to 'dummy' connectors. Volume constraint made restrained cables with covered connectors the selected option. Therefore, connector covers would have to be removed and stowed and the cable untethered prior to making any connection. ISS resource connections to the Core Rack and IRs must be performed by the crew. In order for the Core Rack to control EM(s) inside the IRs; all power, data, and gas connections from rack to rack must be done by the crew on orbit. Any configuration of an active rack isolation system demands that the IRs cannot remain attached to their pivots during operation, but instead must be attached to actuator shafts; which must be installed by the crew. Maintenance operations which require rotation of any one rack demands that either all rack to rack connections be dismantled or that all three racks are rotated out as one unit. A rotation of all three racks as one unit in O g is believed to be the easier of the two options. The racks must be latched together at the top and the IRs removed from actuator shafts and reinstalled upon their pivots as a minimum. Sample insertion and exchange in furnaces demand crew interface with SSFF faceplates, EMs (and associated connectors, tools, hoses, cables), and gloveboxes. Maintenance or upgrade activity depend upon the crew to complete. Some EMs may require command inputs from the crew via the Portable Computer System (PCS).

Approach

A high level task analysis titled SSFF-HETAG-01, "Space Station Furnace Facility Task Analysis-Gross", is provided to systematically identify and characterize major human interface issues and serve as a data source for SSFF-HER-01, "Human Engineering Requirements", test plans and procedures, and SSFF-HECTA-01, "Human Engineering Critical Task Analysis" The Gross Task Analysis gives a beginning point for decision making concerning mock up builds for human subject testing by examining the following subtasks: Lab site checkout, Logistics site checkout, Core and Instrument Rack installation, Core System activation, Nominal EM operations, Facility nominal shutdown, and Maintenance/Upgrade.

Crew interface requirements for SSFF were initially derived in the following manner: NASA-STD-3000, '"Man-Systems Integration Standards" is the original human factors requirements document which covers all aspects of crew-flight concerns; much information that does not apply to a furnace facility. A designer, whether for facility, experiment module. or orbital replaceable unit (ORU) would have to search through extraneous information and attempt to extract that which applied to their particular task. SSP 50005, "ISSA Flight Crew Integration Standards" is derived from NASA-STD-3000 and contains only requirements that pertain to Space Station. This is still not specific enough to help a designer of a furnace facility, an experiment module, or an ORU identify applicable requirements rapidly. Therefore, SSFF-HER-0 1, "Space Station Furnace Facility Human Engineering Requirements" is derived from the above two parent requirements documents and is also influenced by design, specific intent, and the Station Interface Control Document (ICD). SSFF-HER-01 is divided into three sections: Facility, Experiment Module, and ORU sections. A verification matrix is included with each section. This is to simplify the designer's task; i.e., a furnace designer would use the Experiment Module Section of SSFF-HER-01 for human factors requirements applicable to experiment modules.

SSFF-HECTA-01, "Space Station Furnace Facility Critical Task Analysis" (CTA) is an extension and elaboration of the "Gross Task Analysis" (GTA) in matrix format: step by step task description accompanied by applicable SSFF-HER-01 paragraph numbers, crew posture necessary, and notes (tools or not). The Critical Task Analysis can be used as an input to the Crew Procedures development process by systematically examining flight crew tasks

which are critical to mission success. The task analysis begins with the Core Rack and the Instrument Rack installed in the Lab but not connected, proceed through interconnection of racks and ISS, sample loading and unloading, ORU changeout, and Rack rotation. The analysis ends with sample removal after the first increment of operation. This document is continuously updated with the cycle of design change and crew interface tests.

"Space Station Furnace Facility Human Engineering Development Test Plan", SSFF-HEDTP-01, presents tests to be performed with human subjects, program objectives, anticipated results, facility and hardware requirements, reference to CTA sections, lead time required, test dates, and required report dates. The following tests are documented for Phase I: Rack Face Plate Configurations, Avionics Air Coupling, Utility Interface Concepts, ISPR Lowered Floor, Electrical Connector Selection, Single Vs. Multiple Rack Rotation, ORU Changeout, Crew Interface Port Placement and Utilization, Glovebox Installation and Operation with Rack Rotation, Sample Port Location and size. For Phase II design: the following tests are documented: Active Rack Isolation System (ARIS), ORU Changeout. For Phase III design- Alternative ORU placement; ORU Changeouts are repeated, and EM/Glovebox protrusions.

Mock ups

A lot of effort was concentrated on building a mockup of ISS Lab envelope containing SSFF. A rigid material (4 x 8 foot sheets of 3/8 inch thick Styrofoam covered with paper), called Fome Core™ covers the wood frames forming the ceiling and opposite wall of SSFF envelope. A raised plywood floor represents racks that will be in that plane and the station stand off, with 'Z' plates (connects resources from station to racks) made of Fome Core. Three racks (not ISPR quality) are used to represent the Core Rack, IR1 and IR2. The cold plates, flight boxes, and EMs are constructed of Fome Core and installed inside the racks. Most of the cold plates and Gas Distribution System (GDS) mock ups can slide out of the Core Rack along with their contents. Initially, many devices were only Fome Core volumes until more could be learned about the device's exact dimensions. Real connectors, fasteners, and cables are mounted on devices that either have to be installed on orbit, or their failure rate suggests that maintenance operations will be necessary. Cables and hoses are represented by various sized tygon tubing. The degree of fidelity is greater in areas where there is known crew interface.

Many components of SSFF mock up changed during project development. For example, IR1 has contained a Fome Core mock up of a cylindrical furnace with the sample port in the front, a movable carousel, and wooden mock ups of the sample cartridges installed. Beside it was a larger cylindrical furnace with an excess port in the top, and attached side hinges which allowed it to be pulled straight out of the IR1 front and rotated to rest in a vertical position in front of the Core rack. These were replaced later with two rectangular vertical furnace mock ups with both sample ports in the front. IR2 contains two horizontal cradle-mounted furnaces, which are envelope representations only. All the avionics boxes in IR2 are mounted on slides which can be pulled out into the aisle. Two types of gloveboxes were constructed; flexible and rigid. The flexible version was constructed of canvass and Plexiglas. The rigid model was made of Fome Core and Plexiglas.

A mockup of the PCS with a "D" type connector to interface with the crew interface port (CIP) is provided. Velcro is placed on the instrument rack face plates for placement of the lap-top computer with the top of the monitor approximately four and one-half feet from the floor.

The three racks may be fixed together on orbit so that they rotate as one assembly, to avoid disconnections between the Core and the instrument racks. Some critical devices must be located in the back of the racks; either behind furnaces or other equipment. This may make rack rotation necessary several times during the removal/installation of these devices. In order to study this task, each rack is individually counterbalanced with a configuration of pulleys and weights, therefore, all three racks may be locked together at the top and rotated out as one unit, i.e., the whole unit is counterbalanced and will remain in whatever position is needed without any further restraints or attention. Additions or changes inside any rack alters the weight of the rack; counterbalancing is then dealt with by changing the weights in the pulley-weight configuration for that specific rack.

After the initial build, unexpected questions or configurations can be quickly investigated with this mockup which includes surrounding ISS envelope, boxes, lowered floor, shutoff valves, maintenance kill switch, crew interface port, connectors, racks, passthroughs and accesses, pull-out shelves, face plates, some cables and hoses, EMs, Gas Distribution Systems, station 'Z plate plus prototype connectors, and rotating (individual or ganged) counterbalanced racks.

Some crew interface issues had to be addressed before the comprehensive mockup described above was complete. Other tasks examined required the crew to be 'upside down' relative to the local vertical. In these tests, the racks were positioned horizontally with face plates up. Rack to rack interconnections could then be studied with subjects lying on top the rack plates with their heads at the bottom of the racks. This was the crew position most likely to be used for feeding cables through the rack side access panels for rack to rack connections. Mockup hardware was

limited to the area of concern: i.e., the lowered floor, connectors, quick disconnects, tethered cables, labels, and a station 'Z' plate prototype for connecting rack to station resources. In another example, an Avionics Air issue concerning a flex hose connection within the one half inch clearance between the Core Rack and IR1&2 was mocked up with two racks, PVC pipe, flex hose, and a clamp, The two racks were positioned one-half inch apart, side-by-side, and face down in the floor. The rack-to-rack coupling was to be done through the access in the top of the instrument rack and through the side access cut-outs in both the instrument and Core racks. The whole SSFF envelope, detailed components in all three racks, etc. was not necessary to determine that the more space allowed between the solid portion of the air duct and the side access inside IR1, the easier the task of attaching the flex hose between the Core Rack and IR1 duct. Concerns about a box's connectors interfering with tool-hand clearances while securing its flange fasteners was answered with the Fome Core mockup of the box itself complete with real (not flight) connectors, fasteners, and actually using an allen wrench to release a fastener. Understanding the relationship of portions of SSFF configuration was greatly instrumental in building the comprehensive mock up.

Testing with Subjects

Anthropodemetric measuring of many subjects was done prior to testing in an effort to employ the abilities and limitations of the ISS physical dimensions range of 5th percentile (40 year old Japanese female) to 95th percentile (40 year old American male). Seven was the average number of subjects participating pertest and most were working on the SSFF project in various disciplines. Whether individual or group tests, the subjects were briefed as to the hardware description, order of expected tasks, and test objective. The reminder that these tasks must be performed in O g was enforced as much as possible. Where applicable, an 'ergo-chair' was used, which held a person as close as possible to the neutral body position experienced in O g. The test conductor followed a step-by-step checklist of each specific task, which was part of the written test procedure. Comments and suggestions made by the subjects during testing was recorded by the test conductor. Subjects usually completed a questionnaire at the end of a test session. A group test was held for Phase III design - Alternative ORU placement only. For this test, a group of four subjects were allowed to discuss the problem among themselves and physically move shelves and boxes within the racks until they were all satisfied. The physical configuration agreed upon by the group was then recorded. With a mock up in place and an understanding of the required crew task involving the hardware; specific test plans and test procedures for each area investigated were developed.

Evaluation

Methodologies range from pure observation, analysis, questionnaries, to ranking, or some combination of these methods. For example, SSFF is required to maintain the rack face plates in place during operation in order to contain CO_2 for fire suppression. The face plates must also meet the launch load requirements and provide noise attenuation. These requirements complicate crew tasks such as maintenance and servicing by limiting SSFF adjacent rack access, handrail interference, visual, and reach envelopes because the face plates must be opened or removed to perform the tasks mentioned. Although frequent need for the crew to access SSFF racks is not expected, known tasks such as gas bottle change-out prompted the investigation of different face plate configurations through a mock up and test subjects to attempt to arrive at the most 'user-friendly' solution that still allows the face plates to meet more critical requirements. Prior to the mock up demonstration, a preliminary matrix was built with other disciplines participating, listing the requirements of the rack face plates, Goals were associated with each requirement. Symbols indicated each requirement and its intended goal as well as overlapping of unrelated requirements through their goals. Numbers replaced the symbols associated with each requirement based upon a scale of 5 being very important and 1 being hardy important. All values assigned to each requirement were added and then normalized to give every requirement a weight, Four face plate designs were evaluated using the computed weights for the agreed upon requirements. In addition to this described preliminary analysis, seven subjects participated using a test procedure of step-by-step motions outlined for each design configuration. Motion was constrained by the use of an 'ergo chair' to mimic the neutral body posture characteristic of O g. Individuals were timed with a stop watch for the time difference required between the designs as they performed the tasks. The time required to perform the tasks was averaged for each faceplate design. Subjects were asked to complete a questionnaire ranking each design according to effort associated with selected crew interface requirements. Also, their comments were noted on their procedure sheet by the test conductor. The subjects picked the same configuration as the preliminary analysis had indicated to be the 'best design', with the ranking of the other designs varying.

In other instances, the whole range of subject sizes, i.e., 5th percentile (smallest) to 95th percentile (largest); would express the same opinion. This was true with the on-orbit installation testing. IR1 was rotated into the aisle at about the 80' position with its back panel removed. The four flight boxes launched bolted inside the Core Rack were removed and installed into IR1 from the back of the rotated rack. The flange fasteners on the Peltier Pulser and the FSCU (Furnace Signal Conditioning Unit) opposite the operator's position could not be seen or reached. The problem existed for the entire body size range of subjects. Immediate comments such as "I can not see the fasteners," or "I can not reach behind the box." were made. In other instances, no problems were detected, and favorable comments would be common. This was true with the slide out shelves on which the stowed for launch devices were mounted as well as for the Gas Distribution System (GDS) configuration. The quick disconnects used

for the GDS integration testing received no comment because the subjects were able to make the connection. User comments are the best source of ideas since they may reveal why particular errors are occurring. Comments are collected while the user is working, since impressions given after a task is complete are often sketchy. The group testing relied upon their internal discussion and agreement upon device placement, The test reports, which include the evaluations, must be distributed to the designers as shown in Figure 2.

Results

The crew interface organization and flow of work for SSFF is best described in Figure 2. Notable outcomes from this iterative cycle include, but are not limited to:

The fidelity of SSFF hardware features pertinent to the SSFF trainer were identified through the use of the crew interface mock up and testing. More information is available concerning task sequence, completion times, and rack positions (rotated or not). This all feeds directly into crew procedures which will be developed much later in the project. Details such as clearance problems, special tools or modifications concerning ISS provided tools become known. Verification and clarification of many human factors requirements were accomplished as well as configuration feasibility of selected features SSFF which involve human interface. Launch configuration modifications were suggested as a result of problems encountered during on-orbit installation procedures. Misconceptions among various disciplines concerning the physical layout are quickly brought out in tests.

Conclusion

Incorporating crew interface issues early into a design demands an iterative process of design, evaluation, and test with the test results feeding back into design. This can be conceived as slowing down the design process or levying extraneous requirements onto the design. However, building simulated or informal prototypes actually gives a project something tangible for others to see, and stimulates thought and progress. A detailed understanding of the &sign and the necessary crew tasks has enhanced matching the intent of crew interface requirements with the design. Once a comprehensive mockup of SSFF was in place, unexpected questions, design changes, and 'what if scenarios could be tested and produce results quickly. This effort should greatly reduce or negate user related design changes late in the process. Crew time must be used as effectively as possible since it is limited and expensive. Early interest in SSFF design and crew interface issues help identify hardware fidelity and items that will be crucial in the crew trainer. It is the goal of this crew interface endeavor to focus SSFF on the materials research, not 'down time' involving lengthy and difficult crew activities. Figure 2 is a flow chart of the relationship between crew interface requirements, designers, human factors task analysis and development test plan, human factors mockup of SSFF, specific test plans and procedures, test reports, and the feedback into design.

References

NASA-STD-3000, Volume I, Revision A, October 1989, "Man-Systems Intimation Standards" This document contains crew systems integration design considerations, design requirements, and example design solutions for development of crew space systems.

SSP 50005, Revision B, August 1995, "International Space Station Flight Crew Integration Standard" This document provides Flight Crew Interface requirements applicable to the ISS extracted from NASA-STD-3000. The paragraph numbering remains consistent with NASA-STD-3000 and all references to figures and tables not contained in this document are the same figures and tables contained in NASA-STD-3000.

Cash, Martha (1996), SSFF-HER-01, "Space Station Furnace Facility Human Engineering Requirements" This is a requirements document extracted from NASA-STD-3000 and SSP 50005 and is limited to human factors requirements applicable to a furnace facility. The document is separated into three sections; Facility, Experiment Module, and Orbital Replaceable Units (ORUs).

Cash, Martha (1996), SSFF-HECTA-01, "Space Statinn Furnace Facility Critical Task Analysis" The critical task analysis document is a step-by-step description of furnace facility tasks with must be performed by the crew. Each step contains an explanation of how many crew are involved, which paragraph of SSFF-HER-01 applies, and whether or not tools are needed.

Cash, Martha (1996), SSFF-HETAG-01, "Space Station Furnace Facility Task Analysis-Gross." Major crew interface issues are identified and characterized.

Cash, Martha (1996), SSFF-HEDTP-01, "Space Station Furnace Facility Human Engineering Development Test Plan" This document presents tests to be performed with human subjects, program objectives, anticipated results, facility and hardware requirements, reference to CTA sections, lead time required, test dates, and required report dates.

SSP 52007-ICD-SSFF, May 13, 1996, "Space Station Furnace Facility Interface Control Document" This document specifies and controls the design interfaces between the SSFF and the ISS.

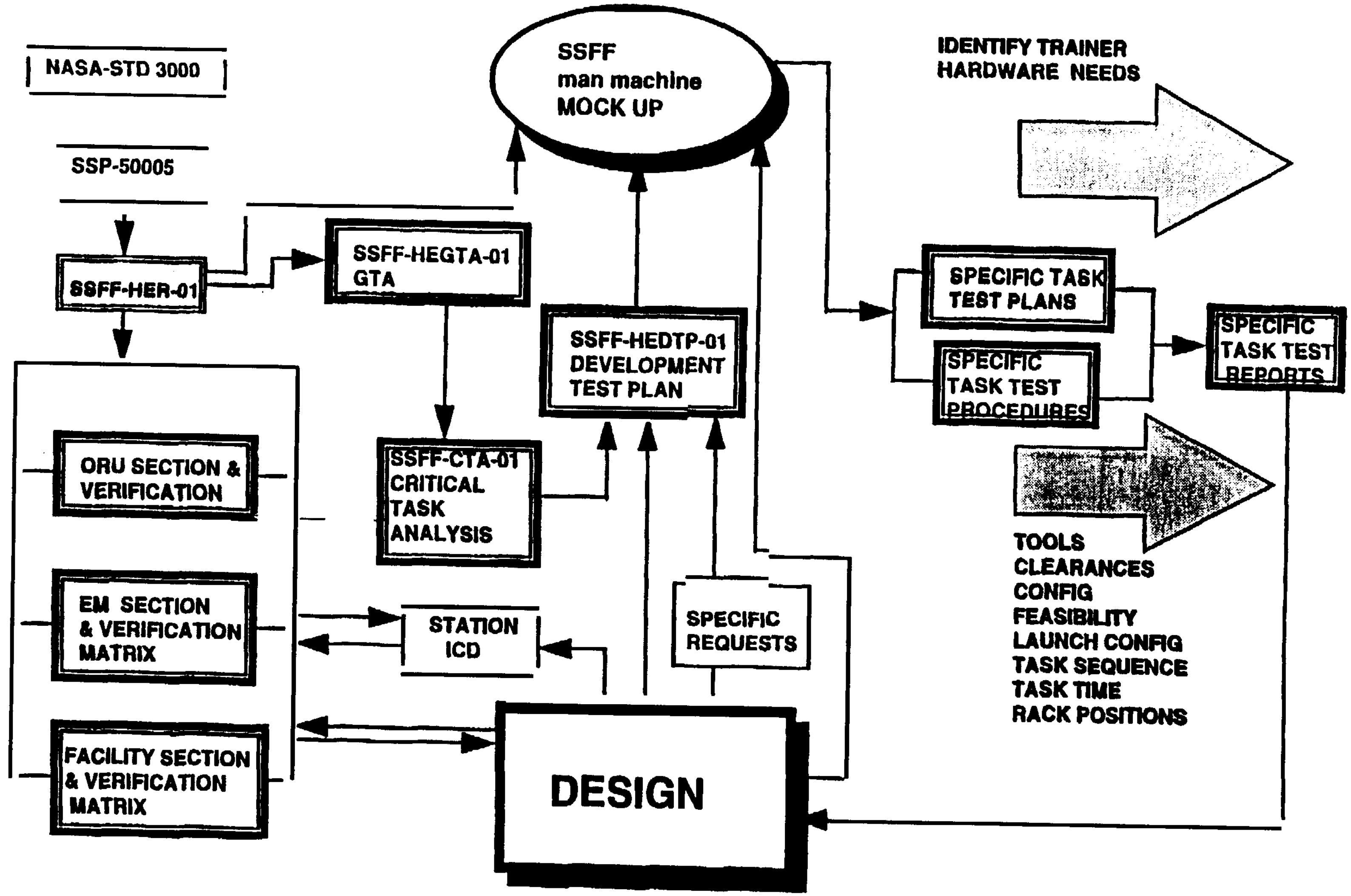


Figure 2. SSFF Man machine Interface Task Flow

URC97026

The MUSES Satellite Team and Multidisciplinary System Engineering

John C. Chen,[†] Alfred R. Paiz,[*] and Donald L. Young[*]

‘Department of Mechanical Engineering
North Carolina A&T State University

5-6/80

*Jet Propulsion Laboratory
California Institute of Technology

Abstract

In a unique partnership between three minority-serving institutions and NASA's Jet Propulsion Laboratory, a new course sequence, including a multidisciplinary capstone design experience, is to be developed and implemented at each of the schools with the ambitious goal of designing, constructing and launching a low-orbit Earth-resources satellite. The three universities involved are North Carolina A&T State University (NCA&T), University of Texas, El Paso (UTEP), and California State University, Los Angeles (CSULA). The schools form a consortium collectively known as MUSES — Minority Universities System Engineering and Satellite. Four aspects of this project make it unique: (1) Including all engineering disciplines in the capstone design course, (2) designing, building and launching an Earth-resources satellite, (3) sustaining the partnership between the three schools to achieve this goal, and (4) implementing systems engineering pedagogy at each of the three schools. This paper will describe the partnership and its goals, the first design of the satellite, the courses developed at NCA&T, and the implementation plan for the course sequence.

Introduction

In efforts to encourage the inclusion of courses which focus on systems engineering as part of the curriculum, and also to increase the number of minority engineers graduating with skills in the design and engineering of complex systems in a team environment, the Jet Propulsion Laboratory (JPL) took the initiative to form a consortium composed of three minority-serving institutions: North Carolina A&T State University, University of Texas, El Paso, and California State University, Los Angeles. The consortium is collectively known as MUSES — Minority Universities System Engineering and Satellite. MUSES has a two-fold mission: (1) the development and implementation of systems engineering courses to teach the fundamentals of the design of large, complex systems, such as a satellite, and (2) the design and construction of a small satellite to be launched into low Earth orbit in approximately three years in order to collect data on planet Earth, which will be analyzed by the Consortium.

In the increasingly competitive world today, government and industry have realized the importance of cohesive, multidisciplinary engineering teams for product design and realization, and problem solving. Yet in our universities, there is currently a dearth of courses and activities which call on students from various engineering disciplines to collaborate in achieving a common design goal. Experience with MUSES shows that a diverse groups of individuals motivated to explore new frontiers can self-organize as a team to complete complex engineering project.

Through the efforts of the MUSES team, minority-serving universities are poised to enrich the pool of future systems engineers produced in the United States. Although MUSES created an exciting design experience for the students, the program's intent was to support the universities' goals related to reinventing the teaching of design engineering. In this regard, the summer workshop featured a self-organized product implementation team strongly coupled to a curriculum development process. The MUSES team sowed the seeds for defining how engineering will be taught and practiced.

Project History

The project began in the Summer of 1995, when faculty representatives from the three universities participated in the Summer Design Institute held July 24 to August 4 at JPL. This extended visit provided opportunities for the faculty to: (1) work with JPL personnel to learn methods of concurrent engineering and project management, (2) observe typical JPL design teams working in JPL's Project Design Center, (3) gain familiarization with system design tools developed at the Project Design Center, (4) discuss systems engineering course elements and implementation, and (5) visit with a faculty member and a student at California Institute of Technology to discuss the satellite design course at their school. The objectives for the summer were to gain familiarity with JPL's Project Design Center, both its architecture and philosophy, and to develop plans for activities in the next academic year (August to May, 1995) and the following summer.

'During the following academic year, plans were made to begin the first Summer Satellite Design Workshop held for ten weeks during Summer 1996 at JPL. Participants in the workshop included faculty members and four students from each of the three schools, and a large contingent of JPL personnel who acted as mentor System Engineers or as mentors on satellite subsystem design. The roles of the students were to act as lead engineers on at least one satellite subsystem each, while contributing as engineering team members on several other subsystems. The faculty members were not involved directly in subsystem design, but were present to mentor the students as needed to solve analytical or other technical problems. A second role the faculty members played was in developing Systems Engineering courses to be implemented back on their respective campuses. The following describes the two activities resulting from the 1996 Summer Satellite Design Workshop: (1) the design of the satellite — called Urania, and (2) the development of the Systems Engineering Curriculum at NCAT.

The MUSES Satellite Team

The students' goal for the 1996 Summer Workshop involved the initial design of a small, low Earth-orbit satellite with a mission to study planet Earth. Realistic constraints on size, power consumption, mass, and cost were specified for the MUSES team. These constraints were stipulated by the requirements that the satellite:

1. be launched as a secondary payload on a Delta II rocket to reduce cost, which essentially limits the dimensions to approximately 30 x 40 x 40 cm;

2. weighs a maximum of 30 kg, again constrained by the secondary payload option;

3. uses flight-tested, off-the-shelf technology to decrease risk, design time, and cost;

4. consume no more power than it can collect using the solar panels to be mounted on its surfaces; and

5. cost no more than $5 million, including all hardware, software and labor, but excluding the launch cost, which was assumed to be donated by external sponsors.

Other aspects of the Summer Workshop included regular meetings with technical personnel from JPL to resolve technical issues and to ascertain progress, and with Human Resources personnel for team building and learning conflict resolution. At the end of the Workshop a formal presentation was made by the team to a panel of technical experts from JPL, and a completed technical proposal was reviewed. The satellite will now be described.

Fig. 1: The 1996 MUSES students, posing with a full-scale mock-up of the Urania satellite.

The initial satellite was designed within the constraints described above. The MUSES team named the satellite Urania, which is one of nine muses of Greek mythology. Urania's overall dimensions are 30 x 40 x 40 cm and it weighs just over 28 kg. Its shape is that of a six-sided cylinder covered on five of the six sides and the top with solar cells. The remaining side is always pointed away from the sun in order to provide a low-temperature surface on which to mount heat radiators for the electronics and instruments. The bottom surface is mounted with the two imaging instruments — one imaging in the near- to mid-infrared for cloud imaging, and one imaging in the visible spectrum for earth resources studies — and thus will always be directed toward earth. The satellite will be stabilized with a 4-m long, gravity-gradient boom which will be deployed after the satellite is released from the launch vehicle into its orbit. The power consumption of the satellite is limited to 17 W maximum, which will require all subsystems, including the instruments, to share this power, and to be placed into "sleep" modes until it is needed. The on-board computer for command and data handling is an A80 186 processor with

600 Mbit DRAM memory for data storage. The satellite has a minimum design life of one year, with a goal of three years.

The mission designed for Urania, as mentioned earlier, is to study planet Earth, and as such its launch is planned to be at an inclination of 42 degrees and at an altitude of 1100 km; this results in a single orbit of 112 min. Three science instruments will be carried as payload: (1) an infrared cloud imager for ocean color mapping and measurements of distributions of water vapor or ice on Earth's surface; (2') a visible-spectrum camera for measuring mineral matter distributions, population densities, and vegetation variations; and (3) a global positioning system (GPS) receiver for the satellite's navigation purposes, and to perform scientific measurements in conjunction with other GPS receivers of other satellites. Science information gained from the satellite provides insight into the geological and environmental phenomena affecting our planet.

The engineering subsystems of the satellite include:

1. Structure/packaging:
2. Mechanisms;
3. Power;
4. Thermal control;
5. Command and data handling-software and hardware;
6. Telecommunications; and
7. Attitude control.

Each of theses subsystems, along with their integration into the satellite, is described in detail elsewhere [1, 2].

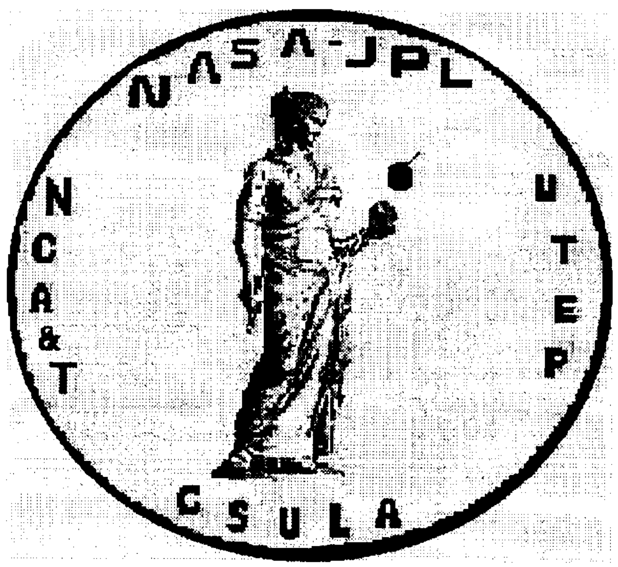


Fig. 2: Logo of the MUSES team, showing the Muse, Urania, holding the Earth as the satellite Urania orbits above.

Multidisciplinary Systems Engineering at NCAT

Due to the fact that the System Engineering curriculum developed is different for each school because of programmatic and administrative constraints. the following is a description of the curriculum to be implemented at NCAT only.

A two-course sequence will be offered in consecutive Fall and Spring semesters of each year. The courses will be open to seniors and graduate students of any engineering major, since the emphasis is on multidisciplinary system design. The first course can be taken as a Technical Elective, while the second will substitute for the Capstone Design course, which is a graduation requirement for the undergraduate curriculum of all engineering majors at NCAT.

Contents of the first course will consist of topics necessary for Systems Design and Engineering but will not be specific to any one system, while the second course will be purely the technical design of a system. The purpose of this is so that the course sequence maybe used to teach the design and engineering of any system, which will most likely change from year to year. Topics of the first course will include team building, technical communications, creative problem solving in multidisciplinary teams, selected issues from total quality management (TQM), and engineering design tools (e.g. software tools). Substantial time will also be devoted to the definition of systems, and analysis of common engineering systems.

The second course in the sequence will focus solely on the design of a single system. In addition to satellite design, other topics may include a solar boat design for a competition sponsored by the American Society of Mechanical Engineers, or the design of experiment modules to be launched with sounding rockets for scientific purposes. The current plan is to implement the courses under the General Engineering division, so that students of any major, in the College of Engineering may take these courses for credit, and to highlight their multidisciplinary nature,

Future Plans

The 1996 Summer Workshop culminated in September with the MUSES team traveling to the Small Satellite Conference sponsored by Utah State University in Logan, Utah. A poster presentation of Urania was made by the team, and favorable comments were received from the attendants.

The immediate plans of the MUSES team is to forge ahead with planned curricular activities for the coming academic year, and to plan for a follow-on Summer Workshop for 1997. It is hoped that several members of the original team will return to refine the initial satellite design. Most importantly, the MUSES team, along with JPL mentors, is seeking to promote the team's activities to private industry and government in hopes of obtaining funding to actually construct the final satellite designed, and to realize its launch, perhaps by 1999. The team's best hope is to compete and win against other universities in future NASA-sponsored satellite design and construction opportunities.

Acknowledgment

We gratefully acknowledge the leadership provided by Ms Bettie White at NASA's Office of Equal Opportunity Programs, Dr. Yvonne Freeman, formerly with NASA, and the Jet Propulsion Laboratory's Minority Science and Engineering Initiatives Office. These programs agreed with our vision of how engineering education can be broadened and improved by the project described in this paper, and provided the necessary funding for its conduct.

We would also like to acknowledge those who participated and are continuing to participate in this venture:

North Carolina A&T State University: Professors Esther Hughes, Eric Cheek, Ganelle Grace, Robert Li, Yong-Duan Song, Shih-Liang Wang, and students Nicole Wright, Kimani Silva, Edgar Masters, Darryl Parker;

University of Texas, El Paso: Professors Mary Clare Robbins, Bryan Usevitch, Scott Starks, and students M iguel Ordaz, Ivan Munoz, Guillermo Velasquez, Joe Riccillo;

California State University, Los Angeles: Professors Kamran Karimlou, Anjan Bhaumik, and students Zoila Herrera, Veronica Lacayo, Luis Rodriguez, Steve Hopkins;

Jet Propulsion Laboratory: Kim Leschly, Cecilia Guiar, Ann Dalton, and other scientists and engineers;

Industrial partners: Mr. Ray Shibata, Hughes Space and Communications Systems, El Segundo, California, and Mr. Richard Croxall, TRW, Redondo 'Beach, California.

References

1 MUSES Program Review Presentation, presented to the Program Review Panel, Jet Propulsion Laboratory, Pasadena, CA (1 996).

2 MUSES Technical Proposal, presented to Jet Propulsion Laboratory, Pasadena, CA (1996).

URC97027

GROWTH OF BULK WIDE BANDGAP SEMICONDUCTOR CRYSTALS AND THEIR POTENTIAL APPLICATIONS

Kuo-Tong Chen, Detang Shi, S. H. Morgan, W. Eugene Collins and Arnold Burger

NASA/Center for Photonic Materials and Devices
Department of Physics, Fisk University
Nashville, Tennessee 37208

S3-76

Abstract

Developments in bulk crystal growth research for electo-optical devices in the Center for Photonic Materials and Devices since its establishment have been reviewed. Purification processes and single crystal growth systems employing physical vapor transport and Bridgman methods were assembled and used to produce high purity and superior quality wide bandgap materials such as heavy metal halides and II-VI compound semiconductors. Comprehensive material characterization techniques have been employed to reveal the optical, electrical and thermodynamic properties of crystals, and the results were used to establish improved material processing procedures. Post-growth treatments such as passivation, oxidation, chemical etching and metal contacting during the X-ray and gamma-ray device fabrication process have also been investigated and low noise threshold with improved energy resolution has been achieved.

I. Introduction

Over the last decades semiconducting compound materials have been recognized and developed for various electro-optical devices applications. Among them, wide bandgap 1 I-VI compound semiconductors have attracted extensive interest and research efforts due to their potential uses as room temperature X-ray and gamma ray detectors[1-7], IR detectors light-emitting devices (LED) in the visible range[9], tunable solid state laser[10], optical limiting and optical communication devices, etc. One of the most essential steps in the the development of a device is the crystal growth process itself. The ability to produce structurally and chemically pure crystals plays a vital role in the development of the practical device, and currently is also an important subject of NASA microgravity materials research program. The growth of single crystals and materials processing under micro gravity environment can prevent gravitational-induced effects, and produce more homogeneous composition and stress-free materials. Materials characterization techniques are also important in providing basic electrical and defect information, and they can be correlated to the starting material and initial crystal growth process in order to optimize the purification and growth conditions. The actual device can then be fabricated and tested to evaluate its performance and potential for practical applications. The Materials Science and Application Group (MSAG) in the Fisk University Center for Photonic Materials and Devices (CPMD) has investigated the fabrication and evaluation of radiation detectors based on wide bandgap II-VI compound semiconductors and heavy metal halides. These materials not onl y have great potential in medical, industrial and environmental applications, but also in space exploration, as X-ray and gamma ray spectrometers and imaging arrays. In this paper we report recent developments in materials processing and characterization and device fabrication.

II. Material Purification and Crystal Growth

Typically, starting materials which may purchased from commercial vendors with nominal purity of 99.9999% or synthesized from pure elements by stoichiometric weight, need to be further purified and/or adjusted to stoichiometric composition before crystal growth and device fabrication. In our laboratory, zone-refining and vacuum sublimation are two standard processes to achieve this goal. Zone-refining was first introduced in 1952 [11], and has been successfully implemented by us to purify elements, such as Se and Te, and heavy metal halides, such as HgI_2, PbI_2 and BiI_3. The characteristic effect of zone-refining is to accumulate impurities at the ends of an ingot, thus leaving pure material in the central section. Figure "1 shows the distribution of impurity concentrations of Mg, Ag, Cu and Cr along a zone-refined ingot of HgI_2[12].

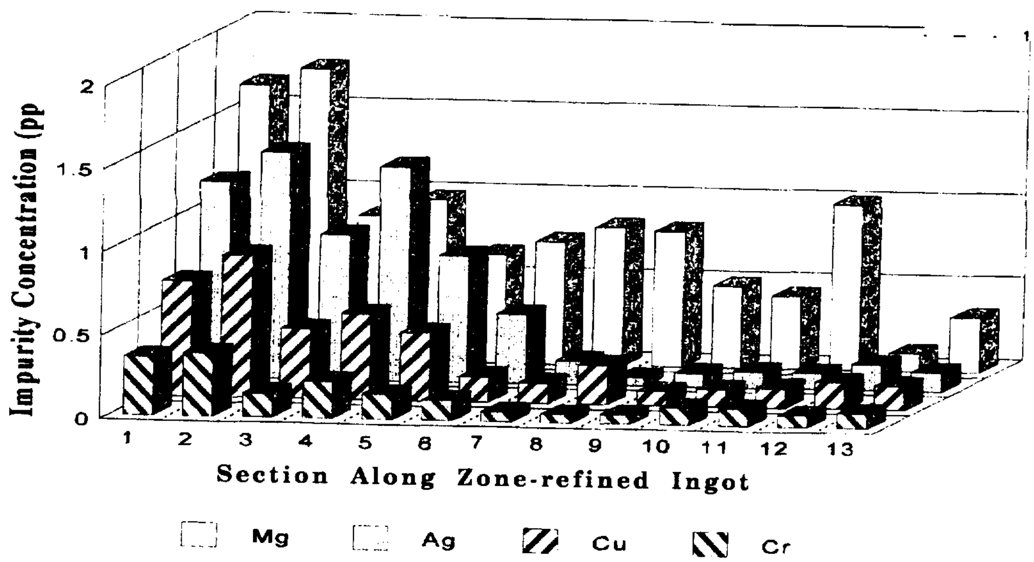
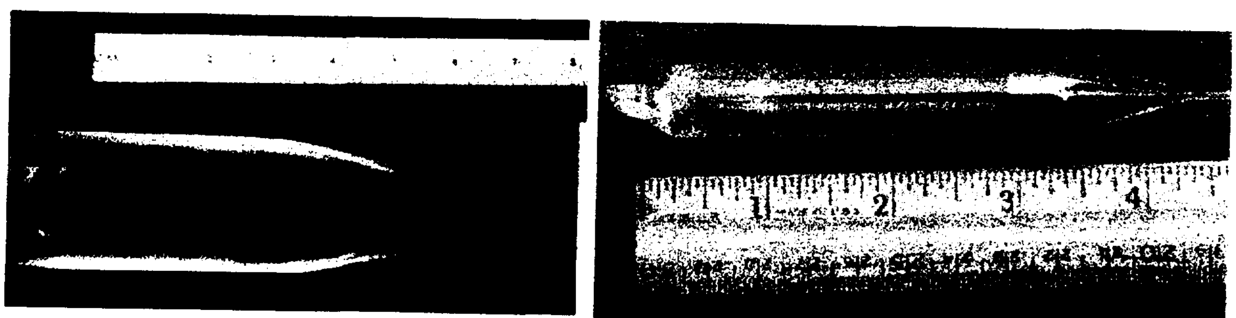


Figure 1. Impurity concentration distribution along zone-refined HgI_2

In this particular case it is clear that most impurities accumulated at one end (1-4) leaving the other section to be the purest part of the ingot. A similar purification effect has been observed for PbI_2[13]. Vacuum sublimation is a routine method used in our laboratory to purify starting materials from impurities having higher (under dynamic vacuum) or lower (closed tube) vapor pressures. Especially, baking the starting material under a dynamic pumping before sealing off the crystal growth ampoule plays an important role in correcting the deviation from stoichiometry of the material.

Two crystal growth methods frequently used in our laboratory are Physical Vapor Transport (PVT) and Bridgman methods. PVT method is the crystal growth under vapor - solid equilibrium conditions. The temperature of the starting material (powder form) is higher than the nucleation/crystal growth region. This imposed temperature gradient leads to a diffusive-convective flow resulting in a net mass transport of vapor species towards the crystal growth site. The vapor species may consist of molecules of the material itself, such as $PbI_2(solid) \rightarrow PbI_2(vapor)$, or dissociated into its separate constituents, such as CdTe (solid) $\rightarrow$ Cd(vapor) + ½ Te_2 (vapor), and residual gases. The reverse process occurs when vapor species nucleate and then continue to condense on the crystal growth interface at a rate of 3-5 mm/day. A typical PVT grown vanadium doped CdSSe single crystal is shown in Figure 2A, currently being investigated for its photorefractive effect and optical limiting properties. Other semiconductor systems, such as HgI_2, II-VI binary and ternary sulfides, selenides and tellurides were also successfully grown from PVT frequently in our laboratory.



grown CdSSe:V

B. Bridgman grown PbI_2

Figure 2. As-grown single crystals in their growth ampoule.

The Bridgman growth method is basically a controlled freezing process taking place under liquid - solid equilibrium conditions. The growth also takes place under a temperature gradient, and the mechanism is to produce a single nucleus from which a single crystal will propagate and grow. This is achieved by allowing the solid - liquid interface to move slowly (5-50 mm/day) until the whole molten charge is solidified. A PbI_2 single crystal is shown in Figure 2B. Compared to other growth methods, Bridgman method is considered to be a rather simple crystal growth method, but several limitations still exist. The Bridgman method can not be applied to a material system which decomposes before it melts, systems having components with high vapor pressure, and materials exhibiting destructive solid - solid phase transformations which will compromise the crystalline quality on cooling the crystal at the end of the growth run. HgI_2 is a typical a material which is not suitable for Bridgman growth. It has high vapor pressure and undergoes a solid - solid phase transition from $\alpha \rightarrow \beta$ at 130 °C before its melting temperature at 258 "C. Therefore HgI_2 can only be grown only by PVT, while PbI_2 lacking such a phase transition can also be grown by Bridgman method.

III. Materials Characterization

Several characterization methods based on spectroscopes, microcopies and thermal analysis techniques have also been employed to investigate and reveal the optical, electrical, structural and thermodynamic properties of the grown crystals.

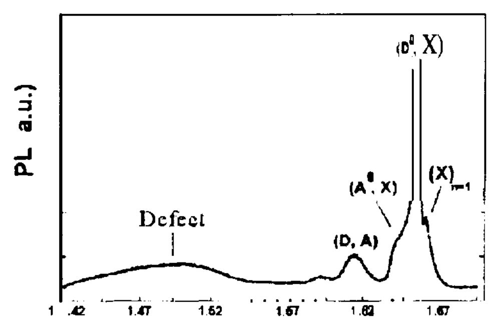
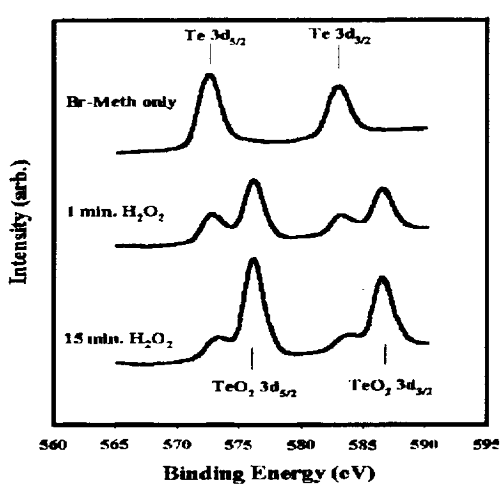


Figure 3. Low temperature photoluminescence spectrum of $Cd_{0.9}Zn_{0.1}Te$.

A. Spectroscopic Characterization - Low temperature (11 K) photoluminescence, x-ray photoelectron spectroscopy (XPS), UV-Vis, infrared spectroscopy (JR) are frequently used, helpful techniques which reveal information of both bulk and surface electronic and optical properties of material. Figure 3 shows a low temperature (11 K) photoluminescence spectrum of $Cd_{0.9}Zn_{0.1}Te$ etched by 5% bromine in methanol and 2% bromine-20% lactic acid in ethylene glycol[14]. The spectrum reveals that the etched surface has a stoichiometric composition and superior structure as evidenced by the appearance of a free exciton in the ground state, $(X)_{n=1}$ and the large $I(D^{\circ},X)/I_{defect}$ (I_{defect} is the broad peak center around 1.5 eV) ratio. Such information is important in understanding surface modifications of a material, and directly relates to the device performance. XPS is a chemical analysis technique which can revealed the composition and chemical species on the surface of materials. Figure 4 shows the XPS spectra of Te 3d peaks of $Cd_{0.9}Zn_{0.1}Te$ etched by H_2O_2[15]. The spectra revealed that H_2O_2 did form a oxide layer on the crystal surface consisting of TeO_2 , the oxidation increasing with H_2O_2 etching time. After a 15 min treatment, the oxide layer thickness could be estimated to be 20-30 Å, based on the relative intensity of Te and TeO_2 peaks. This information is particularly important in the device fabrication since it was found by us to lower the electronic noise of radiation detectors.



B. Microscopic Characterization - Surface morphology can be revealed directly from optical microscopy and high resolution atomic force microscopy (AFM). The morphology of as-grown and

Figure 4. XPS spectra of Te and TeO_2 of H_2O_2 etched $Cd_{0.9}Zn_{0.1}Te$.

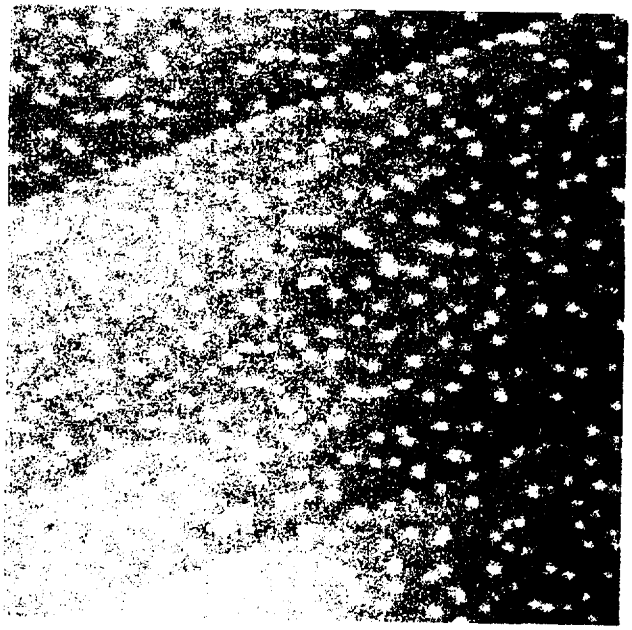


Figure 5 AFM image of cleaved ZnSe crystal, scan size is 500 X 500 nm

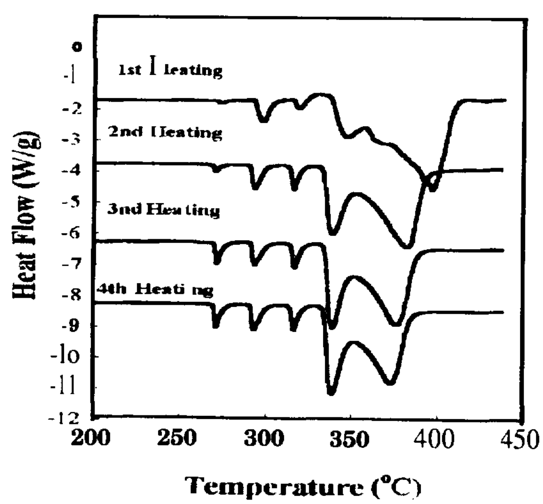


Figure 6. DSC thermograms of Bridgman grown BiI_3.

cleaved crystal surfaces can give important information related to the initial purification process and crystal growth conditions. Figure 5 shows an AFM image of a cleaved ZnSe crystal. This image revealed that second phase precipitation has occurred and that the concentration of precipitates is higher along low angle (approx. 0.1 °tilt) grain boundaries. The density of precipitates is calculated to be about 1013 m^{-2} with 20 nm as an upper limit of the size of precipitate' Using existing phase diagrams, these precipitates have been identified by differential scanning calorimetry (DSC) as a Se-rich second phase existing in ZnSe solid solution. This result was correlated to the initial purification process and crystal growth conditions, and used to improve the growth process of ZnSe crystals with near-stoichiometric composition,

C. Thermal Analysis - Differential scanning calorimetry (DSC), thermomechanical analysis (TMA) and themogravimetric analysis (TGA) have been employed for characterization of mechanical and thermodynamic properties such as phase transitions and melting temperatures, heat of fusion and impurity analysis of materials. Figure 6 shows the DSC thermograms of a Bridgman grown BiI_3 crystal. While the melting temperature for stoichiometric BiI_3 is 408 °C , the Bridgman grown crystals show a depression of melting temperature and some second phases toward low temperature during the first heating run. These second phases were more developed during the second, third and fourth heating cycles, and at the same time the melting temperature has been depressed even more. These second phases were identified from temperature-composition phase diagrams as Bi-rich phases coexisting in a BiI_3 solid solution. PVT grown BiI_3 was also investigated by DSC, not shown here, and the thermograms showed no second phase with a melting temperature near 408 "C. These DSC results clearly show that PVT is the most suitable method for crystal growth of BiI_3 and the melting process produced Bi-rich second phases.

IV. **Device Fabrication and Evaluation**

Room temperature radiation detectors have been fabricated and tested in terms of resistivity and photoresponse. A typical wafer size for radiation detector is about 10×10×2 mm. After wafers are cut from the as-grown crystal, they are mechanically polished, chemically etching to remove surface damage layers and metal electric contacts are applied. Additional surface treatments such as passivation and oxidation are also employed. Each fabrication step may affect the detector performance and therefore unwanted surface effects, such as deviation from stoichiometry, microscopic damage and creation of charge carrier traps, are of great concern. Figure 7 shows the results from current -voltage measurements on the same $Cd_{0.9}Zn_{0.1}Te$ waffer, etched by various etchants[14]. The triangle curve which shows data for the same treatment as in Figure 3, illustrated the most linear ohmicity and the lowest surface leakage current, 5 times lower than other etchants (diamonds: 5% bromine in methanol; circles: 2°/0 bromine plus 20% lactic acid in ethylene glycol.

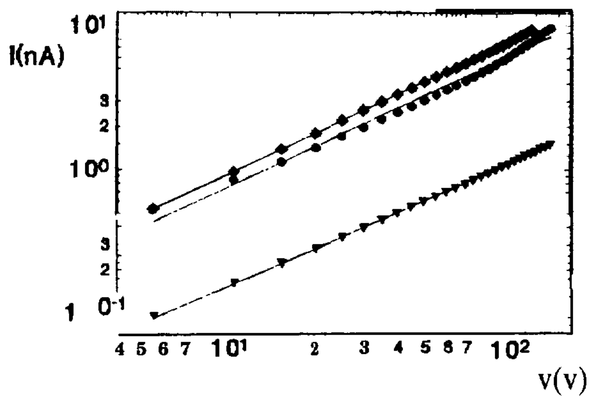


Figure 7. Comparison of I-V curves of $Cd_{0.9}Zn_{0.1}Te$ detector treated by different etchant s.

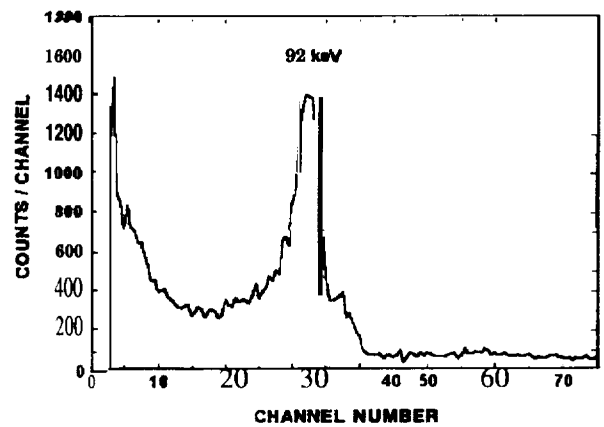


Figure 8. ^{133}Ba spectrum of $Cd_{0.9}Zn_{0.1}Te$.

The good surface quality revealed by the PL spectrum correlates directly to the I-V results in Figure 7. The improvement in the surface leakage current leads to a reduction of the noise threshold allowing for a better low energy x- and gamma-ray spectral response. Figure 8 shows the radiation spectrum obtained from the same sample as in Figures 3 and 7. The noise threshold is the lowest compared to spectra fromsamples treated with the other two etchants, allowing a good energy resolution of 12% of Full-Width-at-Half-Maximum (FWHM) for the 32 keV energy line.

V. Summary

Wide bandgap semiconductor single crystals, such as heavy metal halides and II-W compounds, have been grown by physical vapor transport and Bridgman methods. Zone-refining and vacuum sublimation techniques were used to purify and adjust the stoichiometric composition of the starting material, and were proven to be effective. Several spectroscopic, microscopic and thermodynamic analytical techniques were employed to investigate the optical, electrical and structural properties of crystals. These results revealed information regarding micro- and macroscopic defects, impurities and modifications resulting from source material, growth process, post-growth treatment and device fabrication. Crystal growth and processing conditions have been correlated with this information and were optimized to achieve the purest and highest quality materials for practical device applications. Room temperature x- and gamma-ray detector have been developed and their performance in terms of surface leakage current evaluated and improved. Surface passivation, chemical etching and metal contacting are currently under investigation to further improve device performance, Future works will involve optimization of material purification and crystal growth processes to produce high purity and low defect crystals, development of sensitive material characterization tools allowing a better understanding of defects formation and their correlation with processing conditions.

Acknowledgments

The authors wish to acknowledge the contribution over the years of several of our Fisk graduate students (D. C. Davis, H. Jayatirtha, L. Salary, Y. Zhang, S. U. Egarievwe, S. Lecointe, J. Tong, S. Davis, M. Davis, T. Jounigan and M. Hicks), the helpful interactions with the other researchers at the CPMD (Drs. E. Silberman, M. A. George, H. C hen, and Z. Pan), and external collaborators: C.-H. Su, Y.-G, Sha, M. Volz, and S. L. Lehoczky (NASA/M SFC), R. B. James (Sandia Labs), C. M. Stable, L. M. Bartlett, (NASA/GSFC), T. E. Schlesinger (CMU), D. Nason and van den Berg (EG&G/EM Inc), and L.A. Boatner (ORNL). The financial support from NASA through the Fisk CPMD (Grant # NAGW-2925) as well as the additional support from DOE, ONR, BMDO, Wright Lab and NASA/LeRC are fully appreciated.

References

1. N. L. Skinner, C. Ortale, M. M. Schieber and L. van den Berg, *J. Cryst. Growth*, 89, 86 (1 988),

2. A. Burger, W. Shi, E. Silberman, L. Franks, and W. F. Schnepple, *Nucl. Instr. Methods Phys. Res. A, 283,232* (1 989).

3. A. Burger, D. Nason, L. van dern Berg, and M. Schieber, Ch. 3, Volume 43 of Semiconductors and Semimetals series, "*Semiconductors for Room Temperature* Nuclear Detector Applications", editors: T. E. Schlesinger and R. B. James, Academic Press., "New York (1995).

4. D. C. Dominique, R. B. James, H. Feemster, R. Anderson, A J. Antolak, D. H. Morse, A. E. Pontau, H. Jayatirtha, A. Burger, X. J. Bao, T. E. Schlesinger, G. S. Bench, and D. W. Heikkien, *Mat, Res. Symp. Proc.*, 302,335 (1 993).

5. J. C. Lund, F. Olschner, and A. Burger, Ch. 3, Volume 43 of Semiconductors and Semimetals Series, ''*Semiconductors for Room Temperature Nuclear Detector Applications?'*, **editors**: T. E. Schlesinger and R. B. James, Academic Press., New York (1995).

6. S. U. Egarievwe, L. Salary, K.-T. Chen, A. Burger, and R. B. James, in "*Gamma-Ray Detector Physics and Applications*", Vol. 2305, edited by E. Aprile, 167, SPIE, Bellingham, WA (1 994).

7. S. U. Egarievwe, K.-T. Chen, A. Burger, and R. B. James, and C. M. Lisse, J. X-Ray Sci. And Technol., 6(1 996).

8. A. A. Khan, W. P. Alfred, B. Dean, S. Hooper, J. E. Hawkey, and C. J. Johnson, *J. Electro. Mater., 15,* 181 *(1986).*

9. M. A. Haase, H. Cheng, J. M. De Puydt, and J. E. Potts, J, *Appl. Phys.*, 67, 448(1990).

10. R. H. Page, K. I. Schaffers, L. D. DeLoach, G. D. Wilke, F. D. Pate], J. B. Tassano, S. A. Payne, W. F. Krupke, K.-T. Chen, and A. Burger, *IEEE J. Quantum Electron.*, accepted for publication (1996).

11. W. G. Pfann. *Trans. AIME, 194, 747 (1952).*

12. K.-T. Chen, L. Salary, A. Burger, E. Soria, A. Antolak, and R. B. James, *Nucl. Inst. and Meths. In Phys. Res. A*, accepted for publication (1996).

13. T. E. Schlesinger, R. B. James, M. Schieber, J. Toney, J. M. Van Scyoc, L. Salary, H, Hermon, J. Lund, A. Burger, 'K.-T. Chen, E. Cross, E. Soria, K. Shah, M. Squillante, H. Yoon, and M. Goorsky, *Nucl. inst. And Meths. In Phys. Res. A,* accepted for publication (1996).

14. H. Chen, J. Tong, Z. Hu, D. T. Shi, G. H. Wu, K.-T. Chen, M. A. George, W. E. Collins, A. Burger, R. B. James, C. M. Stable, L. M. Bartlett, *J. Appl. Phys., 80 (5),* September (1 996).

15. K.-T. Chen, D. T. Shi, H. Chen, B. Granderson, M. A. George, W. E. Collins and A. Burger and R. B. James, submitted for publication, *J. Vat. Sci. Technol.* (1996).

16. K.-T. Chen, M. A. George, Y. Zhang, A. Burger, C.-H. Su, Y.-G. Sha, D. C. Gillies, S. L. Lehoczky.

-3/63

URC97028

Application Genetic Algorithms to an Adaptive Fuzzy Logic Controller

Yee-Ming Chen[a] and Tzer-Zong Wang[b]
[a]Department of System Engineering, Chung Cheng Institute of Technology, Taoyuan, Taiwan
[b]Department of Mechanical Engineering, National Central I.University, Taiwan, R.O.C

L Introduction

Fuzzy control has been proven to be useful approach to emulating the human operator with the expertise to control a plant instinctively without reference to established control theory. The human expert describes his control actions by linguistic rules which can be implemented mathematically by use of fuzzy logic. Unfortunately, the mapping of these rules into set theory is not formalised and arbitrary choices for the membership functions have to be made. Hence, the automatic acquisition of rules is an important objective worthy of serious consideration.

This paper investigates the use of genetic algorithm in the design and implementation of fuzzy logic controllers. Previously, generation of membership functions had been a task mainly done either iteratively, by trial-and-error, or by human experts[1]. A task such as this is an autonomous candidate for a GA since GA's attempt to create membership functions that will cause the FLC to perform optimally. Previous work using GA has focused on the development of rule sets or high performance membership functions; however, the membership functions and rule sets are interdependence, using a hand-designed rule set with a GA designed membership functions or hand-designed membership functions with a GA designed rule set does not use the GA to its full advantage[2,3].Thus, the use of GA's to determine both membership functions and rule sets simultaneously for a near-optimal controller would be a more appropriate methodology.

In this paper, a GA has been use to find both membership functions as well as rule sets. A brief review of the GA is first presented. To show the versatility of the methodology, the simulation result show that the FLC for the cart pole problem can achieve a good performance.

2. A review of a genetic algorithm based fuzzy logic controller

2.1 The genetic algorithm paradigm

At first a standard form of the genetic algorithm[4] is summarized, which is a general-purpose stochastic optimization method for search problems, The parameters values processed by the algorithm are a set (population) of strings which represent multiple points in a search space. The string is a binary figure with a finite length where each bit is called an allele and is decoded by a designer to obtain the objective function value of an individual point in a search space. This function value, which should be maximized by this algorithm, is then converted to a fitness value, which determines the probability of the individual undergoing transitional operators.

The transitional operators change the population of the strings to that of offspring: a next generation. The total number of strings included in a population is kept unchanged through generations. The operators are analogous to the biological terms of crossover and mutation. Before applying these operators, a pair of strings is selected among the population with probabilities proportional to their fitness values(must be nonegative). The selection process, an analogy of natural selection, is conducted by spinning a simulated roulette wheel whose slots have different sizes

proportional to the fitness values of the individuals.

Crossover happens for the selected pair with a crossover probability P_c At first, a crossover site (abit position) is randomly selected, and the strings are crossed and separated at the site. This process makes two new strings, each of which takes after both parents. Finally, mutation happens on these two strings with a probability P_m, which inverts a randomly chosen bit on a string. This selection and operation are repeated until the offspring take over the entire population. The next generation is thus made of offspring of three types: mutated after crossover, crossed over but not mutated , and neither crossed over nor mutated, but just selected.

2.2 The GA as used to evolve FLC structure

The basic idea of FLC centers around the labeling process, in which the reading of a sensor is translated into a label as done by human expert controllers. With experts supplied membership functions for labels, sensor readings can be fuzzified and through fuzzy logic eventually defuzzified to generate analog control commands. In our application, we employ a Gaussian distribution function concerned as a kind of membership function with the input x is expressed by,

$$\mu_{A_i}(x) = \exp[\frac{-1}{2}(\frac{x-c_i}{a_i})^2] \tag{1}$$

where c_i is the central position, and a_i is the width(standard deviation). The two parameters are encoded into the gene. In the cart problem, the state of the problem is described by the two error variables e_θ *and@@*, the angle and the angular velocity of the pole respectively, measured from the vertical. In this simulation, the five fuzzy sets employed are NM(negative medium)., NS(negative small), ZR(zero),PS(positive small), PM(positive medium). The rule set, then, contains twenty-five(5×5) rules to account for every possible combination of input fuzzy sets. The rules are of the form, *IF*(e_θ is [NM,NS,ZR,PS or PM] and $\dot{e}_\theta$ is [NM,NS,ZR,PS or PM]) *THEN* ([*output*]), where *output is one* of the fuzzy sets used to partition the output space. The two input spaces use a total of ten Gaussian distribution, so the string to represent a given rule set and membership function combination would have forty-five alleles(25+20). We will consider a control rule table as a genotype with alleles that are fuzzy set indictors over the output domain. The front section of genotype is formed from the control rule table by going rowwise and producing a string of numbers from the code set. In the latter section for addressing both position and width of membership functions. Again, a concatenated, mapped, unsigned binary coding is appropriate for representing the membership functions as bit strings.

3. Simulation results

In this experiment, GA finds the membership functions as well as fuzzy rules for control of the inverted pendulum. The structure of the inverted pendulum system is composed of a rigid pole and a cart onto which the pole is hinged. The cart moves either right or left, depending on the force exerted on the cart. The control goal is to balance the pole starting from non-zero conditions by supplying the appropriate forces to the cart.

If we let $x_1(t) = \theta(t)$ and $x_2(t) = \dot{\theta}(t)$, then this system can be defined by the following differential equations[5]:

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = \frac{g\sin(x_1) + \cos(x_1)[\frac{-F - mlx_2^2\sin(x_1)}{m_c + m}]}{l[\frac{4}{3} - \frac{m\cos^2(x_1)}{m_c + m}]} \tag{2}$$

where g is 9,8 m/s^2, mass of cart m_c is 1.0 kg, mass of pole *m is 0.1* kg, half length of pole l is 0.5 m, and force F is the applied force in newtons. The objective of the GA in this experiment is to find the optimal set of membership functions and fuzzy rules for a FLC. We carry out experiments with the following parameters:

1. Population size: 150,
2. Generation size: 600,
3. Probability of crossover: $P_c = 0.7$,
4. Probability of mutation: $P_m = 0.006$,
5. Selection procedure: stochastic universal sampling,
6. Input signals of FLC: two error variables,0 e_θ and e_θ, angle of pole(degrees), angular velocity of pole(degrees/see),
7. Output control signal of FLC: force(Newtons),
8. Type of membership function: Gaussian distribution shaped(requires parameters $a_i c_i$)
9. Fuzzy partitions: NM,NS,ZR,PS and PM,
10. Number of fuzzy rules: 25,
11. Sampling rate: 0.02 see/sample,
12. The constraints on the dynamic system variables: $\pm$10N for F, $\pm$ 20 deg for e_θ,and $\pm$ 50 deg/sec for $\dot{e}_\theta$

For the design problem, we want to minimize the total amount **error,** $E = \kappa_1 \cdot e_\theta + \kappa_2 \cdot \dot{e}_\theta$. where parameters κ_1 and κ_2 are the weighting constants. Therefore, a smaller E represents a higher fitness(GA maximizes performance). Taking the merit of the genetic algorithm, we define the fitness function by

$$Fitness = 1/E \tag{3}$$

This fitness function offers a means for evaluating the performance of the inverted pendulum. For a given set of membership functions, three different sets of initial conditions were considered in the evaluation of each string. The GA begins by randomly generating a population of 360 strings. After 500 generations of evolution by using $\kappa_1 \cdot \kappa_2 = 0.5$, some of simulation results are illustrated in Fig. 12. The three sets of initial conditions were chosen to ensure that the FLC could posses a good spectrum of system states. This is a testimony to robustness of the technique because a GA has improved the performance of the controller for a situation about which it had no prior information.

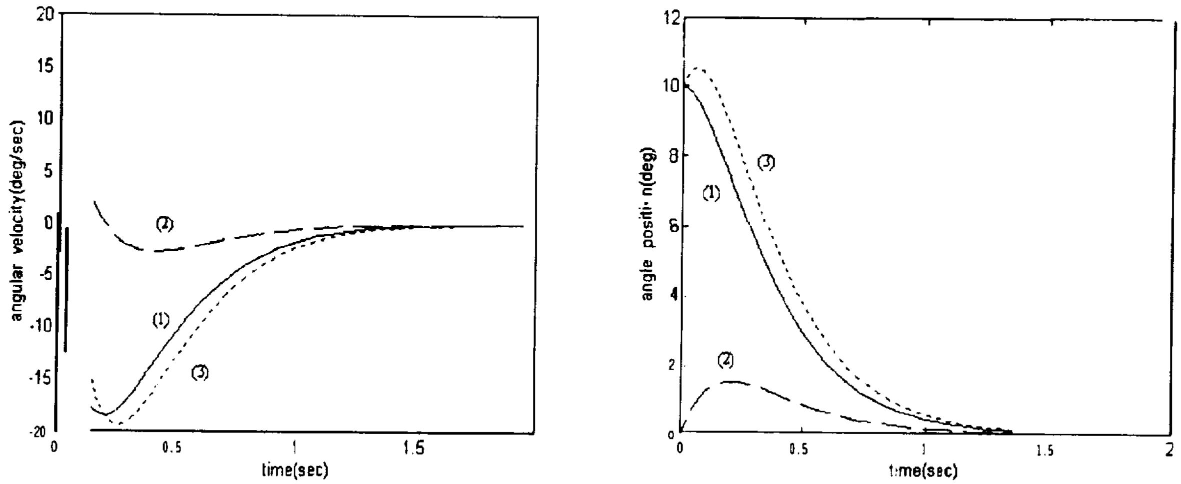
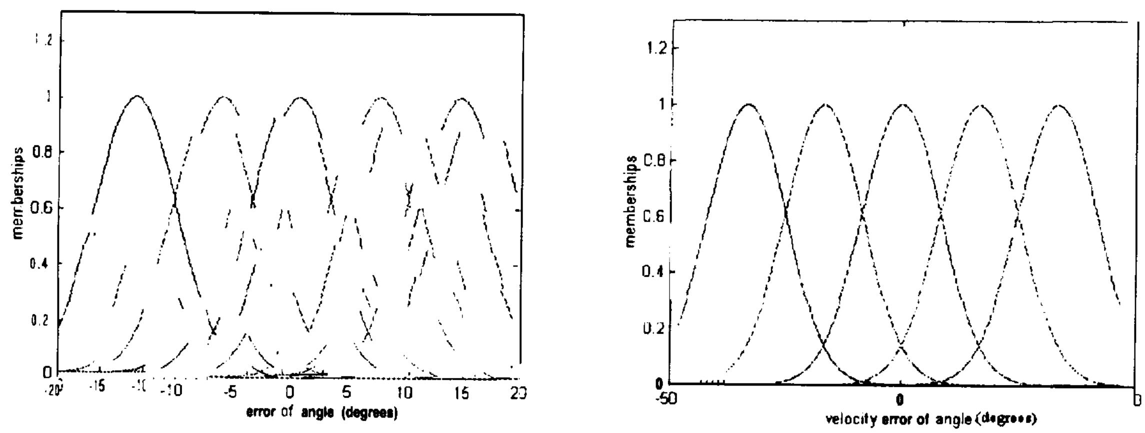
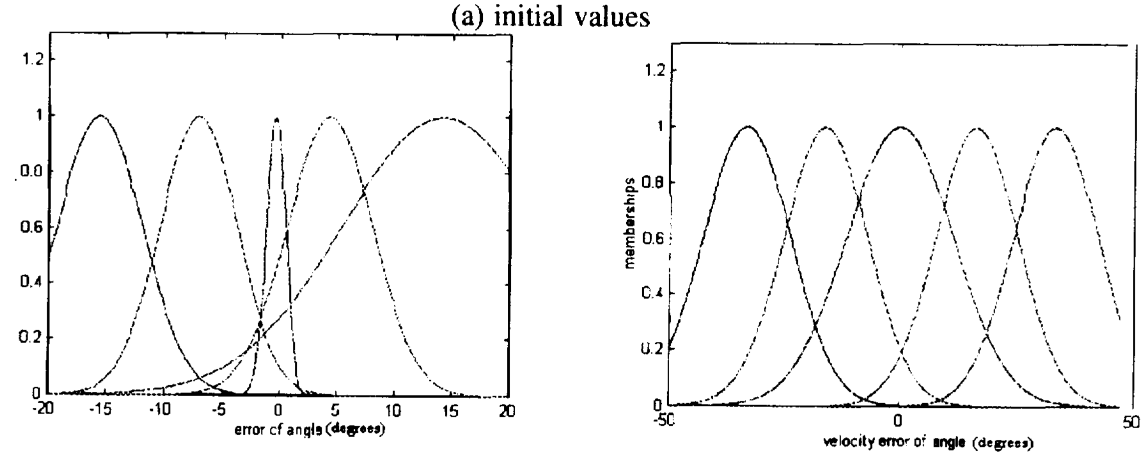


Fig. 1 System responses with differential initial conditions:
(position, velocity): 1:(10,0); 2:(0,20); 3:(10,20)



(a) initial values



(b)final values
Fig. 2 Optimized fuzzy membership functions process

Simulation results also shown in Fig. 2 indicated that the use of GA. Fig 2(b) illustrated the FLC's frizzy sets taken 500 generations in searching the optimal solution. The proposed method show its power in the adaptation process

4. Conclusion

' This paper has shown that the potential for using genetic algorithm to develop an optimal fuzzy logic controller. The simulation results show that the proposed method is effective and efficient. This technique can save time when compared to conventional trial-and-error design procedure. The methodology allows the complete design of both major components of FLC, the rule sets and membership functions, leading to high performance controller which based on the direction of the proposed fitness function.

Acknowledgments

This work was supported by the National Science Control of the "Republic of China.

References

[1] C.C. Lee, " Fuzzy Logic in Control System: Fuzzy Logic Controller, part I, and II" *IEEE trans. on Systems, Man and Cybernetics.* Vol 20, No. 2, 1990, pp. 404-435.

[2] C .L. Karr ," Design of an adaptive fuzzy logic controller using a genetic algorithm" in *Proc. the Fourth Int. Conf. Genetic Algorithms, 1991,* pp. 450-457.

[3] P. Thrift," Fuzzy logic synthesis with genetic algorithms", in *Proc. Fourth Int. Conf. Genetic Algorithm, 1991,* pp.509-5 *13.*

[4] J. H. Holland, *Adaptation in Natural and Artificial Systems.* Ann Arbor, MI: University of Michigan, 1975.

[5] R.H. Cannon,*Dynamics of Physical Systems,*New York: McGraw Hill, 1967.

Page intentionally left blank

-27/63

FEEDBACK IMPLEMENTATION OF ZERMELO'S OPTIMAL CONTROL BY SUGENO APPROXIMATION

C. Clifton, A. **Homaifar** and M. **Bikdash**
The NASA-Autonomous Control Engineering Center
Department of Electrical Engineering
North Carolina A&T State University
Greensboro, NC 27411

Abstract

This paper proposes an approach to implement optimal control laws of nonlinear systems in real time. Our methodology does not require solving two-point boundary value problems online and may not require it off-line either. The optimal control law is learned using the original Sugeno controller (OSC) from a family of optimal trajectories. We compare the trajectories generated by the OSC and the trajectories yielded by the optimal feedback control law when applied to Zermelo's ship steering problem.

1. Introduction

Optimal control [Bryson, 1996; Kirk, 1970] is one of oldest approaches to control engineering. It has many advantages: (1) State and control constraints can be include explicitly. (2) The cost function to be minimized can be often given a simple intuitively appealing interpretation. (3) Optimal control is a very general methodology applicable to multi-input-multi-output, nonlinear, stochastic, or infinite-dimensional systems. Hence, optimal control theory provides a unified approach to stating and solving very general control problems that are at the same time physically intuitive. Unfortunately, optimal control theory suffers from a major disadvantage; namely, solving optimal control problems is in general compuationally difficult except in very special cases where a closed-form expression of the control law can be obtained. These cases include many nonlinear second-order systems and the celebrated linear quadratic regulator. In general however the necessary conditions have no closed-form solution and are at least as difficult to obtain as to solve a nonlinear two-instant boundary value problem (for the control of a system described by deterministic nonlinear ordinary differential equations. When the plant is stochastic or infinite-dimensional, the numerical difficulties are compounded.)

The absence of simple closed-form solutions and online numerical solutions of the general open-loop control problem means that there is no general feedback implementation (except in the neighborhood of an optimal reference trajectory using the well-known neighboring optimal control [Bryson and Ho; 1975].) The lack of feedback implementation is in our opinion the main reason why interest and research conducted in optimal control has greatly diminished.

On the other band, fuzzy-logic controllers (FLCs) are essentially feedback control laws. While theses controllers can be easily made to incorporate the heuristic knowledge of the control engineer, and this can be an advantage in cases where this is about the only knowledge available, designing a FLC using detailed, mathematical, and exact descriptions of the plant is not very well-understood or practiced.

Clearly, using all available knowledge about the system should in principle yield control laws with superior performance. Hence, we investigate in this paper the possibility of designing fuzzy logic controllers that approximate optimal control laws; from another point of view, we investigate feedback implementation of optimal control laws using fuzzy-logic controlled

To illustrate this approach, we consider the Zermelo's problem; that is, the problem of docking a ship going at constant water speed in minimum time in a region of strong water currents using the heading angle as the control i nput. We obtain a family of open-loop solutions of this problem and use it to train the OSC. The resulting trained engine will then be a feedback implementation of (a least-squares approximation of) Zermelo's optimal control. The Sugeno controllers [Buckley, 1993] are capable of approximating any continuous map within an arbitrary accuracy.

This paper is organized as follows. Section 2 provides the necessary background information on the optimal control of the ship steering problem. Section 3 discusses the training procedure used in designing the Sugeno-type controller from the data obtained from the optimal trajectories. Section 4 discusses the generation of training data and the elimination of angle discontinuity. Finally, section 5 summarizes the used procedure and shows simulation results.

2. Zermelo's Optimal Control Problem

The objective of Zermelo's problem is to find a minimum-time path through a region of position-dependent vector

velocity [Bryson and Ho, 1975]. In this problem, a ship must travel in minimum time through a region of strong currents denoted by the two component vector $v(x)$

$$v_1 = v_1(x_1, x_2) \quad \text{and} \quad v_2 = v_2(x_1, x_2) \tag{1}$$

where (x_1, x_2) represent the position of the ship in rectangular coordinates and (v_1, v_2) are the velocity components in the same coordinate system, and the control u is the steering angle O, or $u = \theta$. The magnitude of the ship's velocity relative to the water, V, is a constant. The problem is to steer the ship in such a way as to minimize the time necessary to go from an arbitrary position x_0 to a specified docking position x_f.

The purpose of the generalized Sugeno controller is to approximate the steering angles needed to generate these minimum time paths as a function of x.

The equations of motion are as follows:

$$\dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} \equiv \begin{bmatrix} v_1 + V\cos\theta \\ v_2 + V\sin\theta \end{bmatrix} = v(x) + V \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \tag{2}$$

where $u = \theta$ is the heading angle of the ship's axis relative to the coordinate axes and is the control signal. The Hamiltonian of the system is:

$$H = \lambda_1(V\cos\theta + v_1) + \lambda_2(V\sin\theta + v_2) + 1 \tag{3}$$

and the Euler-Lagrange equations are $\dot{\lambda}_1 = -\frac{\partial H}{\partial x_1}$, $\dot{\lambda}_2 = -\frac{\partial H}{\partial x_2}$, and O $= \frac{\partial H}{\partial \theta}$ whose solution is $\tan\theta = \frac{\lambda_2}{\lambda_1}$. The optimal trajectories satisfy the boundary conditions $x(t_0)$ and $x(t_f)$ specified. Since the Hamiltonian is not an explicit function of time, H = constant is an integral of the system. Furthermore, since the objective is to minimize time, this constant is O. We have five equations to solve for the unknowns $x(t), \lambda(t)$ fort $\in [0, t_f]$, and for t_f.

Following [Bryson and Ho, 1975], we can simplify the two-point boundary problem by solving for $\dot{\theta}$ to obtain

$$\dot{\theta} = \sin^2\theta \frac{\partial v_2}{\partial x_1} + \sin\theta\cos\theta \left(\frac{\partial v_1}{\partial x_1} - \frac{\partial v_2}{\partial x_2}\right) - \cos^2\theta \frac{\partial v_1}{\partial x_2} \tag{4}$$

Equations (4) and (2) are the necessary conditions satisfied by this new and reduced-ordertwo-point boundary value problem. The four boundary conditions are: $x(t_0)$, and $x(t_f)$ are specified. They are used to solve for $\{x(t), \theta(t)\}$ from t_0 to t_f and for t_f itself. Note that if $v(x)$ were constant, then θ would be a constant. In other words, the minimum time paths are straight lines. If $v(x)$ vanes, it is possible for some of the optimal trajectories to intersect at conjugate points x_c. For these trajectories to be considered optimal solutions, the control, $u^*(t) = \theta^*(t)$, must satisfy the following sufficient conditions

$$\frac{\partial^2 H\ (z''(t),\ \theta^*(t),\ \lambda^*(t), t)}{\partial \theta^2} = \frac{v}{V + v_1\cos\theta + v_2\sin\theta} \tag{5}$$

which is clearly is positive &finite if

$$V > v_1\cos\ \theta + v_2\sin\ \theta \ \text{ or if } \ V > \|v\| = \sqrt{v_1^2 + v_2^2} \tag{6}$$

For further discussion of Zermelo's problem, see [Bryson and Ho, 1975].

3. Approximation using the Sugeno Controller

A generalized Sugeno-type controller [Buckley, 1993] is a fuzzy engine mapping a vector $x = [x_1, x_2, \cdots, x_n]^T \in \Re^n$ into u $\in \Re$ where x is interpreted as being a state vector or a measurement and u as a control action. The inference is of the form:

$$R^k\text{: IF } x_1 \text{ is } A_1^k \text{ and } x_2 \text{ is } A_2^k \text{ and...and } x_n \text{ is } A_n^k \text{ THEN } u = y^k = P_k(x) \tag{7}$$

where x_i is the i^{th} component of the input vector x and is a crisp value, A_i^k specifies which among the fuzzy attributes of x_i is tested by rule k, and P_k is a polynomial in $x_1, x_2, \ldots, x_n$ assigned to u by the k^{th} rule. The rules of the original Sugeno controller (OSC) have the following form

$$R^k : IF\ x_1 \text{ is } A_1^i \text{and } x_2 \text{ is } A_2^i \text{ and...} x_n \text{ is } A_n^i\ THEN\ u = y^k = c_0^k + c_1^k x_1 + +c_n^k x_n \tag{8}$$

where $c_0^k, c_1^k, \ldots, c_n^k$ are the consequence coefficients of the k^{th} fuzzy rule. For further discussion of Sugeno-type

controlled see [Buckley, 1993]. Buckley proved that a Sugeno-type controller can approximate any continuous real-valued function in the output space to any degree of accuracy if: (a) the input fuzzy sets have continuous membership functions and (b) a continuous T-norm is being used in the rule evaluation process. This is the universal approximating property of the Sugeno-type controller

Next, we consider approximating the trajectories of the optimal feedback control law by the original Sugeno controller. To do so, we need to determine the coefficients c_0^k, c_1^k, etc. In this paper, we use subscripts to index vectors and superscripts to identify components within a vector In general, the output $\hat{u}$ for the inputs $x_1,...,x_n$ is obtained by the centroid method of defuzzification.

Let (x^j, u^j) be the j^{th} training input/output pair out of a total of J *pairs.* In this paper, these training data are obtained from the generated optimal trajectories. Then the consequence parameters can be obtained by solving a recursive least squares parameter identification problem [Takagi and Sugeno, *1985]* where we determine the unknown coefficients by minimizing the error index

$$\min_{c^1, c^2, \cdots} \quad J = \sum_{j=1}^{J} (u^j - \hat{u}^j)^2 \tag{9}$$

where u^j is the output of the optimal feedback control law and $\hat{u}^j$ is the output of the Sugeno controller. The necessary conditions satisfied by the solution is $\mathbf{ZC} = \mathrm{U}$ where

$$\mathbf{C} = \begin{vmatrix} c' \\ . \\ \mathbf{c}^K \end{vmatrix}, \mathbf{c}^j = \begin{bmatrix} c_0^j \\ \vdots \\ c_n^j \end{bmatrix} \quad \mathbf{U} = \begin{vmatrix} u' \\ \\ u' \end{vmatrix}, \quad Z = \begin{bmatrix} \beta^1 \otimes X^1 \\ \beta^2 \otimes X^2 \\ \cdots \\ \beta^J \otimes X^J \end{bmatrix} \tag{lo}$$

where $\mathbf{Z}$ is a $J \times K(n+1)$ matrix, where $X^j = [1, x_1^j, \cdots, x_?^j$ and

$$\beta_k^j = \beta_k(x^j), \text{and } \beta^j = \beta(x^j) \tag{11}$$

represent the truth values of the rules evaluated at the vector x^j. The least squares solution for C can be calculated recursively by using the following procedure [Takagi and Sugeno, *1985 and* Ljung and Soderstrom, 1986]. Denote the j^{th} row vector of matrix Z defined in (??) by z_j and the j^{th} row of U by u^j. Then, C can be then computed using the iteration.

$$\mathbf{C}^{(j+1)} = \mathbf{C}^{(j)} + \mathbf{S}^{(j+1)} \cdot z_{j+1}^T \cdot \left(u^{j+1} - z_{j+1} \cdot \mathbf{C}^{(j)}\right) \tag{12}$$

$$\mathbf{S}^{(j+1)} = \mathbf{S}^{(j)} - \frac{\mathbf{S}^{(j)} \cdot z_{j+1}^T \cdot z_{j-1} \cdot \mathbf{S}^{(j)}}{1 + z_{j+1} \cdot \mathbf{S}^{(j)} \cdot z_{j+1}^T} \tag{13}$$

where $\mathbf{S}^{(j)}$ is a square $(n(k+1) \times n(k+1))$ covariance matrix at the j^{th} iteration (i.e., after the j^{th} training pair has been acquired and used), and $\mathbf{C}^{(j)}$ the corresponding coefficient vector Then $\mathbf{C}^{(J)}$ at the final iteration is the least squares solution. The initial estimates, $\mathbf{C}^{(0)}$ and $\mathbf{S}^{(0)}$, are chosen as $\mathbf{C}^{(0)} = \mathrm{O}$ and $\mathrm{S}^{(0)} = \alpha I$ where α is a large number and I is the identity matrix

Note that if rule 1 never fires (i.e. $\beta_1^j = \mathrm{O}$ for all j), Then Z is not full rank and $\mathbf{ZC} = \mathrm{U}$ has no unique least-squares solution. Hence, if a rule never fires for the training data given, this rule should be eliminated to make the solution of the least squares problem unique. Also, this rule will not be applicable or relevant in all trajectories similar to the training data.

When the generalized Sugeno controller is used, the above procedure remains largely unchanged except that X now becomes for the case of n = *2*

$$X = [1, x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1^3, x_1^2x_2, x_1x_2^2, x_2^3, \cdots] \tag{14}$$

and the definition of $\mathbf{c}^k$ is to updated accordingly so that y^k defined in Eq. (7) can be expressed in the form $y^k = X\mathbf{c}^k$.

4. Methodology and Procedure

This section proposes a technique to approximate a feedback implementation of optimal controls. It uses the data generated from the optimal control law to identify the coefficients of the generalized Sugeno controller Here, we do not need to solve the two-point boundary-value problem for a arbitrary but given x_0; we only need to generate a family of optimal solutions of $x(t)$ in which an optimal trajectory reaches the final docking position x_f at some final steering

angle Of, To generate one such trajectory, we use integrate Equations (4) and (2) backwards in time from $x(t_f) = x_f$ (the docking position) and $\theta(t_f) = \theta_f$ for any desired t_f until t = O. Optimal solutions are generated for two cases of $v_1(x)$ and $v_2(x)$.

We consider the simple case where the current velocity varies linearly. The objective is to find the minimum-time path from a certain point x_0 to a docking position at the origin. The velocity components of the currents are the following:

$$v_1(x) = -\frac{V}{h}x_2, \qquad v_2(x) = O \tag{15}$$

We generated 18 trajectories for $t_f = 0.6$ seconds. The trajectories along with their time and O contours are shown in Figures 1-2. The magnitude of the ship's velocity relative to water, V, is chosen to be 10 and h, a constant parameter, is chosen to equal 2. The optimal solutions for this case can be obtained in closed-form [Bryson and Ho, 1975], but our figures are generated by backward integration.

The map $\theta = \theta(x)$ modulo 2π has a discontinuity due to the modulo operation. For example, a training trajectory can start with an initial heading angle of 330°, and the angle increases gradually until it reaches 360° (at which θ becomes O") and end at the final heading angle of 10°. The discontinuity occurs at the transition point of 360°/00. Sugeno Engines encounter difficulties in approximating discontinuous maps.

To eliminate this problem, we use two Sugeno engines to approximate the sine and cosine of the heading angle as a function of the state and then to combine them after the approximation. Hence, we approximate $u_1(x) = \cos\theta$, and $u_2(x) = \sin\theta$ using two Sugeno engines and then we combine the two approximates using the formula $\theta = \arctan\frac{u_2}{u_1}$ for use in Eq. (2).

5. Simulation and Results

Now we summarize the procedure followed in this paper and show and discuss the results.

1. Generate the training data
   - Starting from the docking position x_f and a large t_f, integrate Eqs. (2) and (4) backwards in time from final conditions $x(t_f) = x_f$ and $\theta(t_f) = \theta_f$ till $t = O$. *This* will generate one extremum (actually optimum) trajectory for every choice of θ_f.
   - During an integration record the state $x(t)$ and the control $\theta(t)$ as the input and the output training data. The optimal time-to-go for that state is $t_f - t$.
   - Generate a set of trajectories by choosing a set of final values θ_f that is fine enough as to cover the regions of interest in the state space with enough optimum trajectories. Figures 1 and 2 show the generated optimum trajectories for Case 1 along with the associated time and control contours.
2. Perform the least-squares recursion to obtain the consequence coefficients C. There are two sets of coefficients, one for the. sine, one for the cosine.
3. Generate the testing data set. This is achieved by choosing a set of initial conditions z(O). We considered two testing sets:
   - One set was generated by taking the values of the state at the end of the backward integration conducted in step 1 above. We refer to this as Testing Set 1. If the approximation was perfect, the feedback control law will regenerate the optimum trajectories of step 1.
   - Another set was generated more or less randomly near the edge of the region of interest.
4. Simulate the feedback-controlled ship motion.
5. Consider the two performance measures .
   (a) The approximation error index J defined in Eq. (9).
   (b) How close the trajectories of the feedback-controlled ship matched the trajectories of the optimally-controlled ship.

We used five attributes for each input variable. Therefore, there are 25 possible rides, but six of these rides could be eliminated. The original Sugeno Controller yielded a very good approximation. The optimum trajectories and those generated from the approximate feedback law (using Testing Set 1) are shown in Figure 3. The average error index for the OSC is 7.2901.

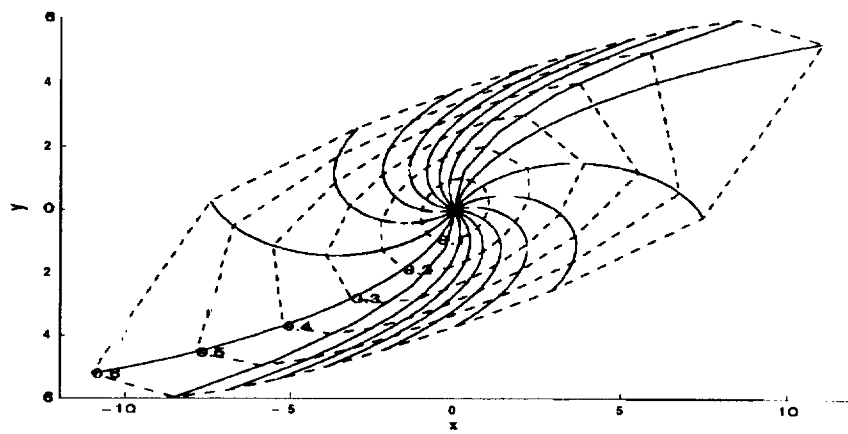


Figure 1: Optimal trajectories with equal-time contours.

6. Conclusions

Sugeno approximation and learning–from–example were shown to yield a powerful and easy-to-use method to implement optimal control in feedback. Since the lack of readily-available feedback implementation is a major limitation of optimal control, this new method is promising and encouraging.

In our approach, optimal control theory is used to generate a set of optimal state and control trajectories usually by backward integration, thus alleviating if not eliminating the need to solve two-point boundary value problems. Next, Sugeno Fuzzy Engines are taught to abstract and approximate the state-to-control mapping from these example trajectories. The original Sugeno engine was used to implement in feedback Zermelo's optimal control.

Acknowledgment

This work is funded by NASA-Autonomous Control Engineering Center under grant number ACE-48146. The authors wish to thank them for their financial support.

References

1. Bryson, A.E., "Optimal Control-1950 to 1985." *IEEE Control Systems.* Vol.16, No. 3, pp. 26-33, 1996.
2. Bryson, A.E. and Ho, Y.C., Applied optimal Control: *Optimization Estimation and Control,* Washington: Hemp-Shire, 1975.
3. Buckley, J, J., "Sugeno-type Controllers Are Universal Controllers. " *Fuzzy Sets & Systems,* Vol. 53, pp.299-303, 1993.
4. Kirk, D.E., *Optimal Control Theory,* Englewood Cliffs, NJ: Prentice Hall, 1970.
5. Lastman, G.J. and Sinha, N.K., *Microcomputer-Based Numerical Methods for Science and Engineering, New York* Saunders, 1989.
6. Ljung, L and Soderstrom, T., *Theory and Practice of Recursive Identification,* Cambridge, MA: MIT, 1986.
7. Sugeno, M. and Kang, G.T., "Fuzzy Modeling and Control of Multilayer Incinerator." *Journal of Fuzzy Sets & Systems,* Vol. 19, pp. 329-346, 1986.
8. Takagi, T. and Sugeno, M., "Fuzzy Identification of Systems and Its Applications to Modeling and Control." *IEEE Trans. on Systems, Man, and Cybernetics,* SMC-Vol. 15, No.. 1, pp. 116-132, 1985.

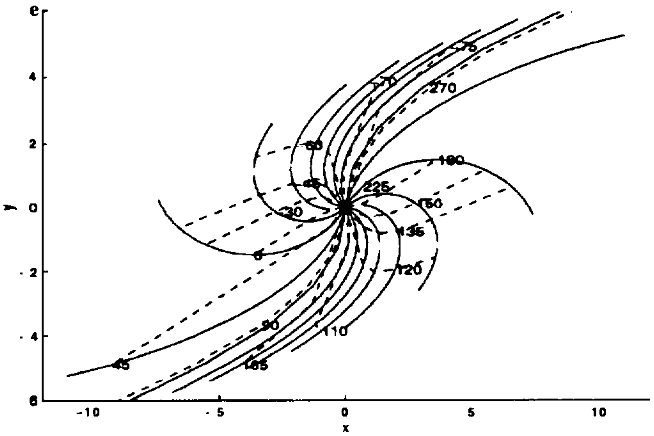


Figure 2: Optimal trajectories with equal-angle contours.

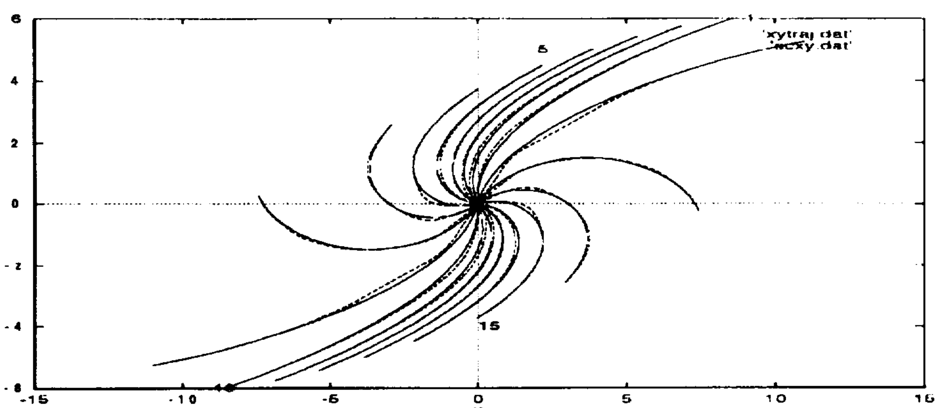


Figure 3: Trajectories generated by the optimal and approximate control laws.

URC97030

INTELLIGENT CONTROL OF FLEXIBLE-JOINT ROBOTIC MANIPULATORS

5-30/63

R. Colbaugh G. Gallegos

Department of Mechanical Engineering
New Mexico State University, Las Cruces, NM 88003 USA

Abstract

This paper considers the trajectory tracking problem for uncertain rigid-link, flexible-joint manipulators, and presents a new intelligent controller as a solution to this problem. The proposed control strategy is simple and computationally efficient, requires little information concerning either the manipulator or actuator/ transmission models, and ensures uniform boundedness of all signals and arbitrarily accurate task-space trajectory tracking.

1. Introduction

The problem of controlling the motion of robotic manipulators in the presence of incomplete information concerning the system model has received considerable attention during the past decade, and much progress has been made in this area. However most of the controllers proposed as solutions to this problem have been designed by neglecting any flexibility associated with the actuator/transmission system and assuming that the actuators are rigidly connected to the manipulator links, As demonstrated in [e.g., 1], joint flexibilities constitute an important component of the complete manipulator dynamic model and thus should be addressed in the controller development process. Recognizing the potential difficulties associated with ignoring the effects of joint flexibility, several researchers have recently considered the problem of controlling rigid-link, flexible-joint (RLFJ) manipulators [e.g., 2-11]. In much of this work, the controller development requires full knowledge of the complex dynamic models for the manipulator and actuator systems [e.g., 2-5]. Research in which controllers are designed with the capability to compensate for uncertainty in the manipulator/actuator system includes adaptive schemes developed using a singular perturbation approach [6,7], which can be used if the joints are sufficiently stiff, and more recent work on robust control strategies and adaptive schemes [8-11] which is valid for arbitrary joint stiffnesses. It is noted that implementation of most of these robust and adaptive controllers requires the calculation of very complex, manipulator-specific quantities, which limits the generality and applicability of these strategies.

This paper introduces a new trajectory tracking controller for uncertain RLFJ manipulators. In contrast with existing schemes, the present strategy is developed using an intelligent control approach which combines ideas from robust control and the recently developed performance- *based adaptive control* methodology [12,13]. This approach effectively exploits the underlying mechanical system structure of the manipulator dynamic model to permit reliance on information regarding this model to be eliminated, and as a consequence overcomes the difficulties associated with previous control methods. Thus the proposed tracking controller possesses a simple and modular structure, is easy to implement, and requires virtually no information regarding either the mechanical or actuator models. It is shown that the controller ensures uniform boundedness of all signals and provides arbitrarily accurate tracking control.

2. Preliminaries

Let $\mathbf{p} \in \Re^m$ **define** the position and orientation of the robot end-effecter relative to a fixed user-defined reference frame and $\theta \in \Re^n$ denote the vector of robot link coordinates. Then the for-ward kinematic and differential kinematic maps between the robot link

coordinates θ and the end-effecter coordinates $\mathbf{p}$ can be written as

$$\mathbf{p} = \mathbf{h}(\theta), \qquad \dot{\mathbf{p}} = J(\theta)\dot{\theta} \tag{1}$$

where $\mathbf{h} : \Re^n \rightarrow \Re^m$ is smooth and $J \in \Re^{m \times n}$ is the end-effector Jacobian matrix.

Observe that there are numerous advantages to formulating the manipulator control problem directly in terms of the end-effector coordinates $\mathbf{p}$. For example, these coordinates are typically more task-relevant than the link coordinates θ, so that developing the controller in terms of $\mathbf{p}$ can lead to improved performance, efficiency, and implementability. If the manipulator is nonredundant (so that m = n) and is in a region of the workspace where J has full rank, then $\mathbf{p}$ and f? are diffeomorphic and this formulation presents no difficulties. A task-space formulation can also be realized if the manipulator is cinematically redundant (so that m < n) by utilizing, for example, the configuration *control* approach [e.g.. 14, 15]. In what follows, we shall consider nonredundant and redundant robots together and introduce a set of n task-space coordinates x obtained by augmenting p with n — m kinematic functions that define some auxiliary task to be performed by the manipulator [14,15]. To retain generality, we shall require only that the kinematic relationship between θ and x is known and smooth and can be written in a form analogous to (1):

$$\mathbf{x} = \mathbf{h}_a(\theta), \qquad \dot{\mathbf{x}} = J_a(\theta)\dot{\theta} \tag{2}$$

where $\mathbf{h}_a : \Re^n \rightarrow \Re^n$ and J_a ● $\Re^{n \cdot n}$. Observe that for x to be a valid task-space coordinate vector the elements of x must be independent in the region of int crest: thus it will be assumed in our development that J_a is of full rank.

Consider an n link RLFJ manipulator with actuator coordinates $\phi \in \Re^n$ and actuator torques $\mathbf{u} \in$ **W.** The task-space dynamic model for this manipulator system is a $4n$th order differential **equation** relating the end-effector coordinates x and the system control input **u:**

$$\mathbf{F} = H\ddot{\mathbf{x}} + V_{cc}\dot{\mathbf{x}} + \mathbf{G}, \qquad \mathbf{T} = J_a^T\mathbf{F} \tag{3a}$$

$$\mathbf{u} = J_m\ddot{\phi} + K_m\mathbf{T} \tag{3b}$$

where $\mathbf{F} \in \Re^n$ is the generalized force associated with *x, H(x)* $\in \Re^{n \times n}$ is the manipulator inertia matrix, V&(x, $\dot{\mathrm{x}}$) $\in \Re^{n \times n}$ quantifies Coriolis and centripetal acceleration effects, G(x) $\in \Re^n$ is the vector of gravity forces, $\mathbf{P} \in \Re^n$ is the vector of forces and moments exert ed by the end-effector on the environment, and J_m , $K_m \in \Re^{n \times n}$ are positive, constant, diagonal matrices which characterize the actuator dynamics. Note that in obtaining the RLFJ manipulator model (3) we have scaled H, V_{cc}, and G by the joint stiffness, introduced the definition T = $\phi - \theta$, and assumed that actuator rotor motion is a pure rotation relative to an inertial frame. It is well known that the rigid-link manipulator dynamics (3a) possesses considerable structure, For example, for any set of generalized coordinates x, the dynamic model terms H, G are bounded functions of x whose time derivatives $\dot{H}$, G are also bounded in x and depend linearly on $\dot{\mathrm{x}}$, the matrix H is symmetric and positive-definite, the matrix V_{cc} *is* bounded in x and depends linearly on x, and the matrices H and V_{cc} are related according to H = $V_{cc} + V_{cc}^T$. Additionally, *V&(x: x) y = V&(x? y*)$\dot{\mathrm{x}}$ ∀y, and if y and y are bounded then *V&(x; y)* is bounded and V_{cc} *(x, y)* grows linearly with x.

In this paper we shall address the trajectory tracking problem. The control objective for tracking is to ensure that the manipulator/actuator system (3) evolves from its initial state to the desired final state along some specified task-space trajectory $\mathbf{x}_d(t) \in \Re^n$ (where $\mathbf{x}_d$ is bounded with bounded derivatives). In what follows, it is assumed that the manipulator/actuator system state $\theta, \dot{\theta}, \phi$, and $\dot{\phi}$ is measurable. Observe that the dynamic

model (3) consists of two cascaded dynamical systems. One consequence of this structure is that the rig-id-link manipulator input **T** cannot be commanded directly, as is assumed in the design of controllers at the link torque input level, and instead must be realized as the output of the actuator' dynamics (3b) through proper specification of the actuator control input **u.** The structure of the RLFJ manipulator dynamics (3) suggests partitioning the control system design problem into two subproblems: regard **T as** the control input for the subsystem (3a) and specify the desired evolution of this variable $\mathbf{T}_d(t)$ in such a way that if $\mathbf{T} = \mathbf{T}_d$ then accurate tracking would be achieved, and then specify the actual control input **u so** that **T** closely tracks $\mathbf{T}_d$. This approach to controller design is adopted in this paper, so that the proposed control system consists of two subsystems: an adaptive strategy that provides the (fictitious) control input $\mathbf{T}_d$ required to ensure that the system (3a) pm-forms as desired, and a robust control scheme that determines the (actual) control input **u** which guarant ees that the system (3b) evolves with **T** closely tracking $\mathbf{T}_d$.

3. Tracking Control Scheme

Let $e = \mathbf{x}_d - \mathbf{x}$ denote the task-space trajectory tracking error and $\mathbf{E} = \mathbf{T}_d - \mathbf{T}$ represent the link torque tracking error. Consider the following tracking controller for RLFJ manipulators:

$$\begin{aligned}
\mathbf{F}_d &= A(t)\ddot{\mathbf{x}}_d + B(t)\dot{\mathbf{x}}_d + \mathbf{f}(t) + k_1\gamma^2\mathbf{w} + k_2\gamma^2\mathbf{e} \\
w &= -2\gamma\mathbf{w} + \gamma^2\dot{\mathbf{e}} \\
\mathbf{T}_d &= J_a^T\mathbf{F}_d \\
\mathbf{u} &= \mathbf{f}_0(t) + [\mathbf{b}(t)\,\mathrm{sat}(\frac{\mathbf{s}}{\epsilon})] + k_a\mathbf{s}
\end{aligned} \tag{4}$$

where the notation $[\mathbf{gh}] = [g_1 h_1, g_2 h_2, \ldots . g_n h_n]^T \in \Re^n$ (for any two n-vectors g, h) and $\mathrm{sat}(g) = [\mathrm{sat}(g_1), \mathrm{sat}(g_2), \ldots . \mathrm{sat}(g_n)]^T \in \Re^n$ (with $\mathrm{sat}(\cdot)$ the standard saturation function) is introduced, $s = \dot{\mathrm{E}} + \lambda\mathrm{E}$ is the weighted torque-torque rate tracking error, $\mathbf{f}_0(\mathbf{t}), \mathbf{b}(t) \in \Re^n$ **are** robust control terms, $\mathbf{f(f)} \in \Re^n$ **and** $A(t)$, $B(t) \in \Re^{n \times n}$ are adaptive gains! and $k_1, k_2, \gamma, k_a, \epsilon, \lambda$ are positive scalar constants. The robust control terms f., b are smooth vector functions which are defined in the proof of the Theorem below, and the adaptive gains $\mathbf{f}, A, B$ are adjusted according to the following simple update laws:

$$\begin{aligned}
\dot{\mathbf{f}} &= -\sigma_1\mathbf{f} + \beta_1\mathbf{q} \\
\dot{A} &= -\sigma_2 A + \beta_2\mathbf{q}\ddot{\mathbf{x}}_d^T \\
\dot{B} &= -\sigma_3 B + \beta_3\mathbf{q}\dot{\mathbf{x}}_d^T
\end{aligned} \tag{5}$$

where $\mathbf{q} = \mathbf{e} + k_2\mathbf{e}/k_1\gamma - \mathbf{w}/\gamma$ is the weighted and filtered position-velocit y tracking error and the σ_i and β_i arc positive scalar adaptation gains.

The stability properties of the proposed tracking strategy (4),(5) are summarized in the following theorem.

Theorem: The control scheme (4),(5) ensures that (3) evolves with all signals (semiglobally) uniformly hounded provided γ is chosen sufficiently large and **b is** properly defined. Moreover, the trajectory tracking error e,& is guaranteed to converge exponentially to a compact set which can be made arbitrarily small.

Proof: Observe first that the actuator dynamics (3b) can be written

$$J_m\dot{\mathbf{s}} = \mathbf{f}_m(\theta, \dot{\theta}, \phi, \dot{\phi}) - \mathbf{u} \tag{6}$$

where $\mathbf{f}_m(\theta, \dot\theta, \phi, \dot\phi)$ is a smooth function obtained through routine manipulation. Applying the control law (4) to the manipulator dynamics (3) yields the closed-loop error dynamics

$$
\begin{aligned}
&H\ddot{\mathbf{e}} + V_{cc}\dot{\mathbf{e}} + k_1\gamma^2\mathbf{w} + k_2\gamma^2\mathbf{e} + \mathbf{\Phi}_f + \Phi_A\ddot{\mathbf{x}}_d + \Phi_B\dot{\mathbf{x}}_d + V_{cd}\dot{\mathbf{e}} - J_a^{-T}\mathbf{E} = O \\
&J_m\dot{\mathbf{s}} + k_a\mathbf{s} + [\mathbf{b}\mathrm{sat}(\frac{\mathbf{s}}{\epsilon})] + \mathbf{f}_0 - \mathbf{f}_m = 0
\end{aligned}
\tag{7}
$$

where $\mathbf{\Phi}_f = \mathbf{f} - \mathbf{G}$, $\Phi_A = A - H$, $\Phi_B = B - V_{cd}$, and the notation $V_{cd} = V_{cc}(\mathbf{x}, \dot{\mathbf{x}}_d)$ is introduced.

Consider the Lyapunov function candidate

$$
\begin{aligned}
V = {} & \frac{1}{2}\dot{\mathbf{e}}^T H\dot{\mathbf{e}} + \frac{1}{2}k_2\gamma^2\mathbf{e}^T\mathbf{e} + \frac{1}{2}k_1\mathbf{w}^T\mathbf{w} + \frac{k_2}{k_1\gamma}\mathbf{e}^T H\dot{\mathbf{e}} - \frac{1}{\gamma}\mathbf{w}^T H\dot{\mathbf{e}} \\
& + \frac{1}{2}\mathbf{s}^T J_m\mathbf{s} + k_a\lambda\mathbf{E}^T\mathbf{E} + \frac{1}{2\beta_1}\mathbf{\Phi}_f^T\mathbf{\Phi}_f + \frac{1}{2}\mathrm{tr}[\frac{1}{\beta_2}\Phi_A\Phi_A^T + \frac{1}{\beta_3}\Phi_B\Phi_B^T]
\end{aligned}
\tag{8}
$$

and note that V is a positive-definite and proper function of the closed-loop system state if γ is chosen sufficiently large. Computing the derivative of (8) along (7) and simplifying permits the following upper bound on $\dot V$ to be established:

$$
\begin{aligned}
\dot V \le {} & -\lambda_m(Q^*)\,\|\mathbf{z}_1\|^2 - k_o\,\|\dot{\mathbf{E}}\|^2 - k_a\lambda^2\,\|\mathbf{E}\|^2 + \frac{3}{\sigma_{min}}\,\|\mathbf{z}_1\|\|\mathbf{E}\| - \frac{\eta_1}{\beta_{min}}\,\|\mathbf{\Psi}\|^2 \\
& + \frac{k_2k_{cc}}{k_1\gamma}\,\|\mathbf{e}\|\|\dot{\mathbf{e}}\|^2 + \frac{k_{cc}}{\gamma}\,\|\mathbf{w}\|\|\dot{\mathbf{e}}\|^2 + \eta_2\epsilon + \frac{\eta_3}{\beta_{min}}
\end{aligned}
\tag{9}
$$

where $\lambda_m(\cdot)$, $\lambda_M(\cdot)$ denote the minimum and maximum eigenvalue of the matrix argument, respectively, k_{cc} satisfies $\|V_{cc}\|_F \le k_{cc}\|\dot{\mathbf{x}}\|\ \forall\mathbf{x}$, k_{cd} is an upper bound on V_{cd}, v_M is an upper bound on $\|\dot{\mathbf{x}}_d\|$, the η_i are positive scalar constants which do not increase as ϵ is decreased and the β_i are increased, $\mathbf{f}_0$ is any (nominal) estimate for $\mathbf{f}_m$ (for example, $\mathbf{f}_0 = O$ can be used), $\mathbf{b}$ is chosen so that $b_i \ge \max[1, (f_{mi} - f_{0i})^2]$ for $i = 1, 2, \ldots, n$, γ is chosen so that $\gamma \ge \max[1, k_2/k_1]$, $\mathbf{z}_1 = [\|\mathbf{e}\|\ \|\dot{\mathbf{e}}\|\ \|\mathbf{w}\|]^T$, $\mathbf{\Psi} = [\|\mathbf{\Phi}_f\|\ \|\Phi_A\|_F\ \|\Phi_B\|_F]^T$, $\beta_{min} = \min(\beta_i)$, $\beta_{max} = \max(\beta_i)$, σ_{min} is the minimum singular value of the matrix J_a (recall that J_a *is* assumed to be nonsingular in the region of interest, so that σ_{min} is nonzero), and

$$
Q^* = \begin{bmatrix}
\frac{k_2^2\gamma}{k_1} & -\frac{k_2}{2k_1\gamma}(k_{cc}v_M + k_{cd}) & 0 \\
-\frac{k_2}{2k_1\gamma}(k_{cc}v_M + k_{cd}) & (\frac{\gamma}{2} - \frac{k_2}{k_1\gamma})\lambda_m(H) - k_{cd} & -\lambda_M(H) - \frac{k_{cd}}{2\gamma} - \frac{k_{cc}v_M}{2\gamma} \\
0 & -\lambda_M(H) - \frac{k_{cd}}{2\gamma} - \frac{k_{cc}v_M}{2\gamma} & k_1\gamma
\end{bmatrix}
$$

(note that Q^* is positive-definite if γ is chosen large enough). Next let $\mathbf{z}_2 = [\|\mathbf{z}_1\|\ \|\mathbf{E}\|\ \|\dot{\mathbf{E}}\|]^T$ and

$$
Q = \begin{bmatrix}
\lambda_m(Q^*) & -3/(2\sigma_{min}) & 0 \\
-3/(2\sigma_{min}) & k_a\lambda^2 & 0 \\
0 & 0 & k_o
\end{bmatrix}
$$

and notice that Q is positive-definite if k_a is chosen large enough. If ϵ is chosen to be inversely proportional to β_{min} and β_{max}/β_{min} is fixed, then there exist positive scalar

constants η_4, η_5 that does not increase as γ and β_{min} increase, and positive scalar constants λ_i independent of γ and β_{min}, such that V and V in *(8)* and (9) can be bounded as

$$\lambda_1 \parallel z' \parallel^2 + \frac{\lambda_2}{\beta_{min}} \parallel \Psi \parallel^2 \leq V \leq \lambda_3 \parallel \mathbf{z}_2 \parallel^2 + \frac{\lambda_4}{\beta_{min}} \parallel \Psi \parallel^2$$

$$\dot{V} \leq -(\lambda_m(Q) - \frac{\eta_4}{\gamma} V^{1/2}) \parallel z' \parallel^2 - \frac{\lambda_5}{\beta_{min}} \parallel \Psi \parallel^2 + *$$

Now choose two scalar constants V_M, V_m *so* that $V_M > V_m \geq V(0)$, and define $c_M = \lambda_m(\mathbf{Q}) \quad \eta_4 V_M^{1/2} / \gamma$; then choose γ large enough so that $c_M > 0$ (this is always possible). Let $\delta = \max(\lambda_3 / c_M, \lambda_4/\lambda_5)$ and choose β_0 so that $\eta_5 \delta / \beta_0 < V_m$ (this is always possible). Then selecting $\beta_{min} \geq \beta_0$ ensures that if $V_m \leq V < V_M$ then $\dot{V} < 0$. This condition together with $V_M > V_m \geq V(0)$ implies that $V(t) \leq V_M \; \forall t$, *so* that $c(t) = \lambda_m(Q) - \eta_4 V^{1/2}(t)/\gamma > c_M > 0 \; \forall t$ and

$$\dot{V} \leq -c_M \parallel \mathbf{z}_2 \parallel^2 - \frac{\lambda_5}{(\beta_0 + \Delta\beta)} \parallel \Psi \parallel^2 + \frac{\eta_5}{(/\sim. + \Delta\beta)}$$

where $\Delta\beta = \beta_{min} - \beta_0$ and it is assumed that β_{min} is chosen so that $\Delta\beta > 0$. The ultimate boundedness results developed in [13,14] are now directly applicable and permit the conclusion that $\parallel \mathbf{z}_2 \parallel, \parallel \Psi \parallel$ are uniformly bounded and that $\parallel \mathbf{z}_2 \parallel, \parallel \Psi \parallel$ converge exponent ially to the" closed balls B_{r_1}, B_{r_2}, respective] y, where "

$$r_1 = \left(\frac{\delta \eta_{14}}{\lambda_1 (\beta_0 + \Delta\beta)} \right)^{1/2}$$

$$r_2 = \left(\frac{\delta \eta_{14}}{\lambda_2} \right)^{1/2}$$

Observe that the radius of the ball to which $\parallel \mathbf{z}_2 \parallel^2$ is guaranteed to converge can be decreased as desired simply by increasing $\Delta\beta$. ■

4. Conclusions

This paper presents a new solution to the motion control problem for uncertain RLFJ manipulators. The proposed control strategy is simple and computationally efficient, requires little information concerning either the manipulator or actuator/transmission models, and ensures uniform b oundedness of all signals and arbitrarily accurate task-space trajectory tracking. Future research will involve the implementation of the controllers for robotic applications in hazardous and unstructured environments.

5. Acknowledgments

The research described in this paper was supported in part by the Army Research Office, the Department of Energy (WERC), and the NASA Center for Autonomous Control Engineering at the University of New Mexico.

6. References

1. Good, M., L. Sweet, and K. Strobel, "Dynamic Models for Control System Design of Integrated Robot and Drive Systems", *ASME Journal of Dynamic Systems, Measurement, and Control,* Vol. 107, 1985, No. 1. pp. 53-59

2. Spong, M., "Modeling and Control of Elastic Joint Robots", *ASME Journal of Dynamic Systems, Measurement, and Control,* Vol. 109, No. 4, 1987, pp. 310-319

3. Khorasani, K., "Nonlinear Feedback Control of Flexible Joint Manipulators". *IEEE Transactions on Automatic Control,* Vol. 35, No. 10.1990, pp. 1145-1149

4. Dawson, D., Z. Qu, J. Carroll, and M. Bridges, "Control of Robot Manipulators in the Presence of Actuator Dynamics", *International Journal of Robotics and Automation.* Vol. 8, *No.* 1, 1993, pp. *13-21*

5. Nicosia, S. and P. Tomei, "Design of Global Tracking Controllers for Flexible-Joint Robots", *Journal of Robotic Systems,* Vol. 10, No. 6. 1993, pp. 835-846

6. Spong, M., "Adaptive Control of Flexible Joint Manipulators", *Systems and Control Letters,* Vol. 13, 1989, pp. 15-21

7. Khorasani, K., "Adaptive Control of Flexible-Joint Robots", *IEEE Transactions on Robotics and Automation,* Vol. 8, No. 2, 1992, pp. 250-267

8. Lozano, R. and B. Brogliato, "Adaptive Control of Robot Manipulators with Flexible Joints", *IEEE Transactions on Automatic Control,* Vol. 37, 'No. 2, 1992, pp. 174-181

9. Dawson, D., Z. Qu, and M. Bridges, "Hybrid Adaptive Control for Tracking of Rigid-Link Flexible-Joint Robots", *IEE Proceedings D,* Vol. 140, No. 3, 1993. pp. 155-1.59

10. Tomei, P., "Tracking Control of Flexible Joint Robots with Uncertain Parameters and Disturbances", *IEEE Transactions on Automatic Control,* Vol. 39, No, 5. 1994, pp. 1067-1072

11. Qu, z., "Input-Output Robust Tracking Control Design for Flexible Joint Robots". *IEEE Transactions on Automatic Control,* Vol. 40, No. 1, 1995, pp. 78-83

12. Colbaugh, R., K. Glass, and H, Seraji, "Performance-Based Adaptive Tracking Control of Robot Manipulators", *Journal of Robotic Systems,* Vol. 12, No. 8, 1995, pp. 517-530

13. Colbaugh, R., H. Seraji, and K. Glass, "Adaptive Compliant Motion Control for Dexterous Manipulators", *International Journal of Robotics Research,* Vol. 1.4, No. 3. 1995, pp. 270-280

14. Seraji, H., "Configuration Control of Redundant Manipulators: Theory and Implement at ion". *IEEE* Transactions on *Robotics and Automation,* Vol. 5, No. 4.1989, pp. 472-490

15. Glass, K., R,. Colbaugh, D. Lim, and H. Seraji, "Real-time Collision Avoidance for Redundant Manipulators", *IEEE Transactions on Robotics and Automation,* Vol. 11, NO. 3, 1995, pp. 448-457

URC97031

The Onset of the Madden-Julian Oscillation within an Aquaplanet Model

"EDWARD COLÓN AND JAMES LINDESAY

The Center for the Study of Terrestrial and Extraterrestrial Atmospheres, Howard University, Washington, D.C. 20059

S31/45

MAX SUAREZ

Laboratory for Atmospheres/Goddard Space Flight Center, Greenbelt, Maryland, 20771

ABSTRACT

A series of numerical experiments using a two-level atmospheric general circulation model (AGCM) were performed for the purpose of investigating the coupling between sea surface temperature (SST) profile and the onset of the Madden-Julian Oscillation (MJO). The AGCM was modified to run as an aquaplane with all seasonal forcing removed (Hayashi and Sumi, 1986). SST distributions based on the New Global Sea-Ice and Sea Surface Temperature (GISST) Data Set for 1903-1994 (Rayner et rd., 1995) were generated then modified to vary the north-south gradient and tropical temperatures. It was found that the MJO signal did not depend on the SST temperature gradients but rather on the absolute temperature of the equatorial region, EOF analysis revealed that the SST distribution which generated the strongest M-JO signal produced a periodic fluctuation in velocity potential at the 250 millibar level with a phase speed of 15 m s and a periodicity of 30 days which falls within the shortest limit of observed oscillations. This distribution also possessed the coolest equatorial SSTs which suggests that increased stability in the atmosphere favors the occurrence of organized MJO propagation.

1. Introduction

Large scale circulation and convection features in the tropical atmosphere have been observed to fluctuate with a certain characteristic time scales and spatial patterns. Two such global scale phenomena are the El Niño/ Southern Oscillation, which has a time scale of roughly 2 to 6 years, and the Madden-Julian Oscillation (MJO) of a corresponding oscillatory period of 30 to 60 days. The latter phenomenum was first detected by Madden and Julian (1 971, 1972) using spectral analysis of daily rawinsonde data for Canton Island in the central tropical Pacific. Their observations yielded a vertically coherent tropospheric oscillation in zonal wind, pressure and temperature fields with a period of 41 to 53 days. Although they do not offer an explanation for its existence, they do observe that the osculation is coherent through all levels of the troposphere and that it possesses characteristics of a circulation cell, which suggests that deep convection is an important driving mechanism. Over the next few years, several researchers reported the existence of a quasi-40 day cloudiness oscillations over the summer monsoon region of East Asia and India and suggested that they may be related to those found by Madden and Julian. Satellite observations of 30- to 60- day fluctuations in tropical cloudiness that tended to propagate eastward into the central Pacific also seemed to be tied to the phenomenum (Weickmann, 1983 and Weickmann et alia, 1985).

A substantial body of research since then has clarified many of the fundamental characteristics associated with this oscillation. The most notable properties of are summarized as follows:

- The oscillation is characterized by global-scale tropical wind and convection anomalies, including a modulation of the Northern Hemisphere and Southern Hemisphere summer monsoon activity (Krishnamurti and Subrahmanyam, 1982)
- The oscillation is not strictly periodic, but has a preferred time scale of about 30 to 60 days.
- Convection and circulation anomalies associated with the oscillation tend to propagate eastward with time with a speed of 10-15 m/s, which is much slower than the phase speed of the equatorial Kelvin waves with the same vertical wavelength (Madden and Julian, 1972).
- in the tropics, 30-60 day zonal wind anomalies in the lower troposphere are out of phase with those in the upper troposphere and can be defined as possessing baroclinic structure(Murakami and Nakazawa, 1985).
- Coherent fluctuations between extratropical circulation anomalies and the tropical 30-60 day oscillations may exist, indicating possible tropical midlatitude interactions in the above time scale(Weickmann et al., 1985).

The origin of the MJO has remained a enigma since its initial detection. In this study we will look at three explanations: wave-CISK mechanism (Lau and Peng, 1987; Hendon, 1988), evaporation-wind feedback (Emmanuel, 1987; Neelin, Held and Cook, 1987) and the super cloud cluster forcing mechanism (Chao and Lin, 1993; Chao, 1995).

2. Theory

Much of tropical meteorology involves the collective interaction of large-scale circulation dynamics and cumulus convection. The candidate most likely responsible for this interaction is conditional instability of the second kind (CISK) (Charney and Eliassen, 1964). This mechanism is characterized by a low-level convergence field associated with a large-scale meteorological system that lifts surface air to a given height where cumulonimbus convection occurs. The latent heat released by this convection forces a large-scale motion whose low-level convergence in turn helps maintain the convection. Underlying all CISK models is the assumption that deep cumulus convection will not, by itself, maintain the low level convergence necessary for the maintenance of clouds. Consequently, some large-scale, low-level convergence is necessary. Preliminary studies of CISK generating mechanisms pointed to the possibility of Ekman pumping as providing the necessary convergence (Charney and Eliassen, 1964). However, near the equator the term associated with momentum flux divergence vanishes, making Ekman pumping an unlikely candidate for providing the necessary forcing. It is now generally accepted that equatorially trapped internal waves(gravity, Kelvin, mixed gravity-Rossby, and Rossby), which are highly convergent, can produce CISK without the need of Ekman pumping (Lindzen, 1974).

Hendon (1988), used a two level model and a CISK-type cumulus heating scheme to obtain a reasonable simulation of the MJO and identified its structure as that of a nonlinearly coupled Rossby-Kelvin mode. It was found that an east ward propagating CISK mode, consisting of a Rossby wave component to the west and a Kelvin wave component to the east, is in agreement with observations. The Rossby component is continuously generated as the disturbance moves eastward. Consequently, we should be able to observe a meridional wind perturbation accompanying the eastward moving oscillation; this, in fact, has been observed (Madden, 1986) and modelled (Hayashi and Sumi, 1986). This CISK mode is similar in structure to the quasi-steady response (Gill, 1980) to slow eastward moving diabatic forcing. There is a pronounced east-west asymmetry in the equatorial region. Kelvin waves carry the information rapidly eastward, thereby creating eaterly low-level winds in that region. These winds provide inflow to the region of heating and subsequently initiate a Walker-type circulation with rising over the source region and sinking to the east. Additionally, heating would generate Rossby waves which would carry information westward.

We can gain a better understanding of the coupled Rossby-Kelvin mode by examining the equatorially trapped free oscillations using the beta-plane approximation (Hendon, 1988) from which the following arguments have been derived. Asumming that there is no initial meridional forcing in the tropics, we can express the horizontal momentum equations as

$$\frac{\partial u}{\partial t} - \beta y v = -\frac{\partial \Phi}{\partial x}, \qquad (1)$$

$$\beta y u = -\frac{\partial \Phi}{\partial y}, \qquad (2)$$

$$\frac{\partial \Phi}{\partial t} + c^2 (\nabla \cdot U) \quad K^2 (\nabla \cdot U) \qquad (3)$$

where $\beta = 2\Omega/R$, u is the perturbation zonal wind, v is the perturbation meridional wind, Φ is the geopotential, c is the velocity of propagation, U is the sum over perturbations in zonal and meridional wind, and K is cumulus heating factor. For this particular model $c^2 (=gH) = (p/p_o)^{R/C_p} C_p\ 0$ and K? = $(p/p_o)^{R/C_p} C_p\ Q'$ where θ is the potential temperature and Q' is the ratio of combined steady diabatic heating and cumulus heating to $\nabla \cdot U$. The equatorially trapped wave solutions (Holton, 1992) can be derived from Eqs. (l), (2), and (3) so that

$$\begin{pmatrix} u \\ v \\ \Phi \end{pmatrix} = \begin{bmatrix} u(y) \\ v(y) \\ \Phi(y) \end{bmatrix} H_n(y) \exp[i(kx - \omega t)], \tag{4}$$

and the dispersion relation is given by

$$\omega = -\frac{kC}{2n+1} [1 - (K/c)^2]^{1/2}, \tag{5}$$

where n is the meridional mode number, $H_n(y)$ represents the Hermite polynomial of degree n, and k is the zonal wavenumber. The Kelvin wave dispersion is obtained for n = -1 and n = 1 yields the most prominent Rossby wave.

For $K^2 = 0$ (no cumulus heating) the solutions yield a dry Kelvin wave which propagates eastward at $c = ((p/p_0)^{R/C_p} C_p \theta)^{1/2} \approx 60$ m s^{-1} which is roughly three times faster than the induced n = 1 Rossby wave. For $0 \lessapprox K \lessapprox c$ (stable cumulus heating), moisture modified waves are slowed by a factor of

$$[1 - (K/c)^2]^{1/2} = [1 - (Q'/\theta)]^{1/2}. \tag{6}$$

It is obvious that eastward propagating Kelvin waves can be reduced to any phase speed with the appropriate choice of Q'. For $K^2 > c^2$ (unstable condition), the dispersion relation takes the form $i\omega$ where ω is defined by Eq. (5). In this case, all waves are stationary and increasing in amplitude due to the fact that they possess imaginary phase speeds. It is possible that a propagating coupled mode may exist in the linear model when $\theta < Q'$ which implies that the individual Rossby and Kelvin modes will have imaginary phase speed. However, a horizontally coupled propagating mode could exist in the presence of positive cumulus heating.

An alternative mechanism for MJO generation has been proposed by Emmanuel (1987), and Neelin, Held, and Cook (1987) that acts through a feedback between evaporation and surface winds. The major limitation of evaporation-wind (E-W) feedback as a modulating mechanism for Kelvin wave propagation is that it assumes that lower tropospheric easterly equatorial flow is a mandatory basic state when in fact observations have clearly revealed westerlies in the Indian Ocean (Wang, 1988). Additionally, Chao (1994, 1995) has linked the development of the MJO to the cloud cluster teleinduction mechanism. Satellite observations have shown that convective regions associated with this oscillation consist of one or more super cloud clusters and within each of them individual cloud clusters arise, move westward, and then decay within 2 to 3 days. New cloud clusters appear to the east of the existing cloud clusters. The reason that the eastern side of an existing cloud cluster is more favorable than the western side for the new cloud cluster formation has to do with basic flow in the boundary layer being easterly which is also strengthened by the circulation induced by the existing cloud cluster. In other words, the east side provides the primary moisture supply Once the new cloud cluster emerges, it competes for moisture supply with the existing cloud cluster. The existing cloud cluster weakens as it propagates westward into the area of depleted moisture which is consumed by the newly developed cloud cluster. Successive generation of new cloud clusters in the east and the subsequent decay of existing ones give rise to an eastward moving envelope, which is the super cloud cluster associated with the MJO.

3. Model

A two-level general circulation model (Zephyr) based on an older primitive equations model developed by Held and Suarez (1978) is employed to examine MJO development and propagation. It utilizes DYCORE, the standard C-grid (4° x 5°), dynamical core used by NASA/Goddard GCMs. Zephyr is characterized by simple cumulus parameterization, and full a hydrologic cycle. In a manner similar to Hayashi and Sumi (1986), Zephyr is modified to run as an aquaplanet in which earth's surface is covered by a uniform global ocean. This change can be made since the oscillation is primarily dependent on the moisture content of the tropical region and the oscillatory structure of the phenomenum is more readily observable without the effects of intervening land-masses. Zephyr is an atmosphere-only model (AGCM) so that a fixed zonally symmetric sea surface temperature (SST) distribution is employed. The temperature profile is based on the New Global Sea-Ice and Sea Surface Temperature (GISST) Data Set for 1903-1994 (Rayner et al., 1995). GISST is composed of monthly averaged SSTs which have been resolved using eigenvector (EOF) reconstruction methods. Data resolution is 1° x 1° for absolute SST readings back to 1982, and 2° x 2° for anomalous SSTs taken back to 1949. Prior to 1949

anomalous SSTs have a resolution of 5°x 5°. The advent of remote sensing and systematic data collection has allowed for better grid resolution since 1982.

A 4°× 5° zonal average was taken for GISST data set so that it could be utilized by Zephyr. A slightly asymmetric SST profile resulted between the northern and southern hemispheres. Since we wanted to achieve a fairly realistic aquaplanet profile, it was decided to "flip" the southern hemisphere SST readings into the northern hemisphere deleting any anomalous effects due to land-masses and thus providing a truly zonally symmetric SST profile with a maximum temperature of 28.6°C at the equator and a minimum of -2°C at the poles. Since Zephyr was originally intended to examine interannual phenomena dependent on seasonal fluctuations, we had to modify several subroutines in order to observe daily perturbations with the seasonal variability essentially "turned off." Additionally, Topography was flattened in order to remove any orographic forcings and albedo was set at 0.07 which corresponds to that of ocean water. Horizontal diffusion is handled by means of an eighth order Shapiro filter used to curtail *nonlinear* computational instability initiated by the cascade of variance to short horizontal scales.

4. Preliminary Results

An initial experiment was conducted with the goal of obtaining a realistic simulation of MJO-like propagations within the equatorial region of the model. In order to investigate the connection between SST profiles and the MJO, five different distribution with incrementally decreasing pole to equator temperatures were tested. The first run was performed using the original GISST profile. Once the model climatology was deemed acceptable, analysis was made of the 250- and 750-millibar level zonal wind (u), precipitation(p), latent heat flux (FL), velocity potential (χ) and divergence of wind velocity($\nabla \cdot U$). Hovmöller diagrams of u, ϱ, and χ revealed little in terms of a stable eastward propagating oscillatory signal with the characteristic 30-60 day period. However, a weakly defined easterly flow was observed at the equator of order 10 m s-1. The initial GISST profile was modified so that the SSTs were linearly decremented in the tropics yet held constant at the poles. All other parameterizations were held constant and each model was run for 365 days plus an additional 100 days spin up time. Five meridional temperature profiles were examined starting with the initial pole to equator distribution ranging from -2°C to 28.6 °C, and ending with a distribution that ranged from -2°C to 23.6°C. Of the five distributions, the -2°C to 23.6°C profile yielded the strongest MJO signal. The following Hovmöller diagrams and figures are for the this case. Figures 1 and 2 **reveal the** periodic structure of the MJO phenomenum

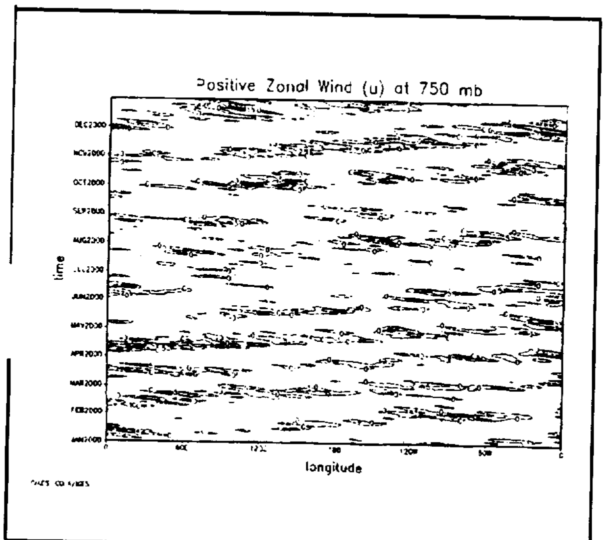


Figure 1. Hovmöller diagram of zonal wind at 750 mb. (0-5 m s^{-1} contour)

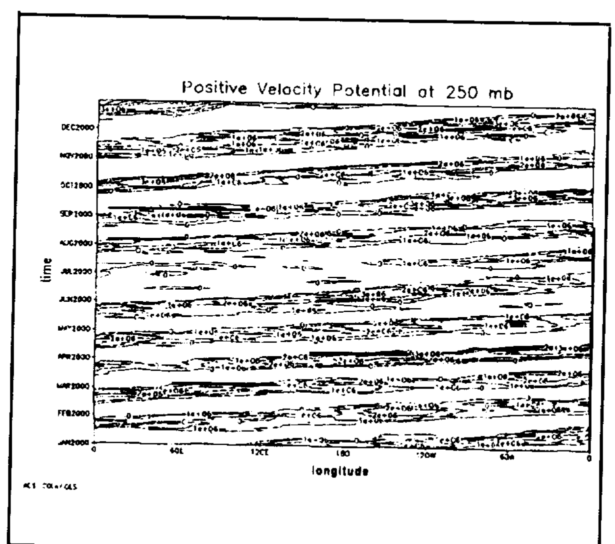


Figure 2. Hovmöller diagram of χ_{250}.

Based on the relatively systematic time evolution of the 30-60 day χ_{250} anomalies, we elected to use an empirical orthogonal function (EOF) analysis of this data to determine the time variation and phase of the oscillation. Figure 3 below depicts the first two eigenvectors (spatial patterns), while figure 4 illustrates the principal component (time series). It is interesting to note that first two principal components describe 40% and 38% of the χ_{250} 30-60 day variance

respectively. Additionally, there is a 900 phase shift between the first and second eigenvalues and principal components. A period of roughly 30 days is observed in the time series which entails a MJO propagation velocity of 15 m S-1.

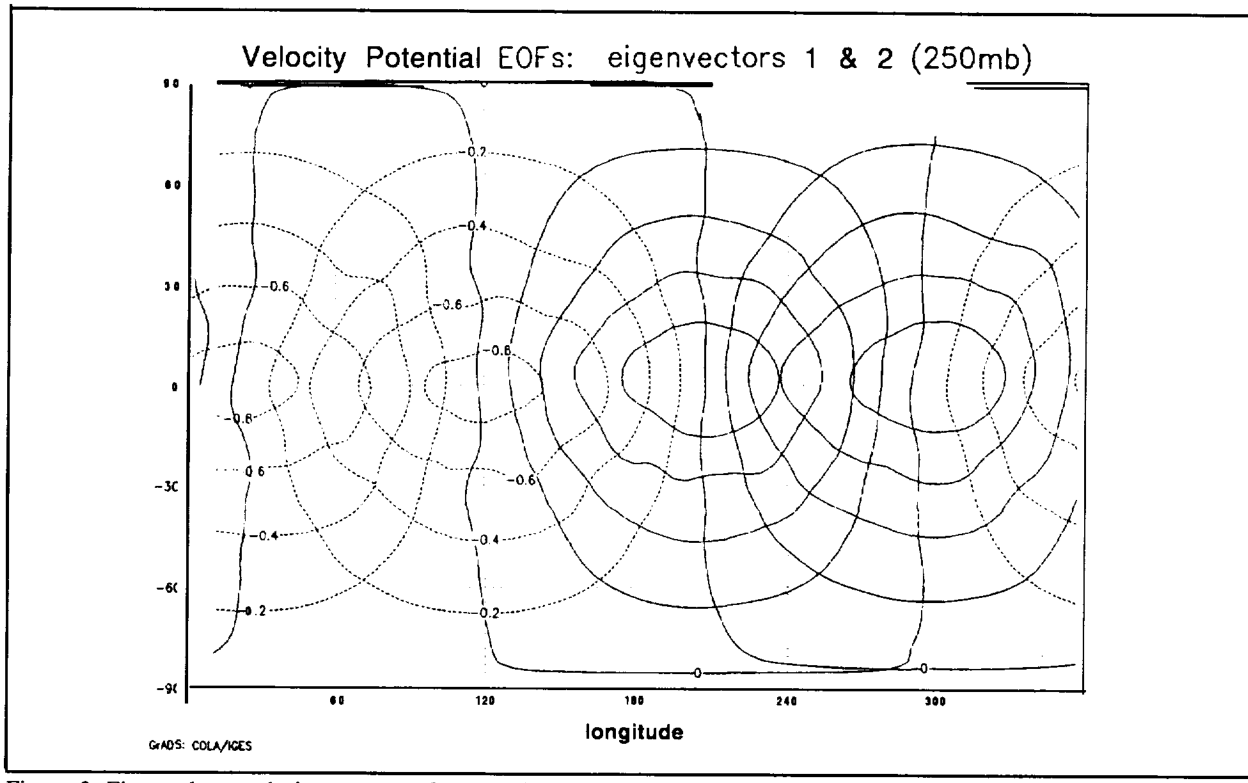


Figure 3. First and second eigenvectors of χ_{250}. Solid line designates positive EOFS, dashed lines designates negative EOFS. Contour internal is set at 0.2.

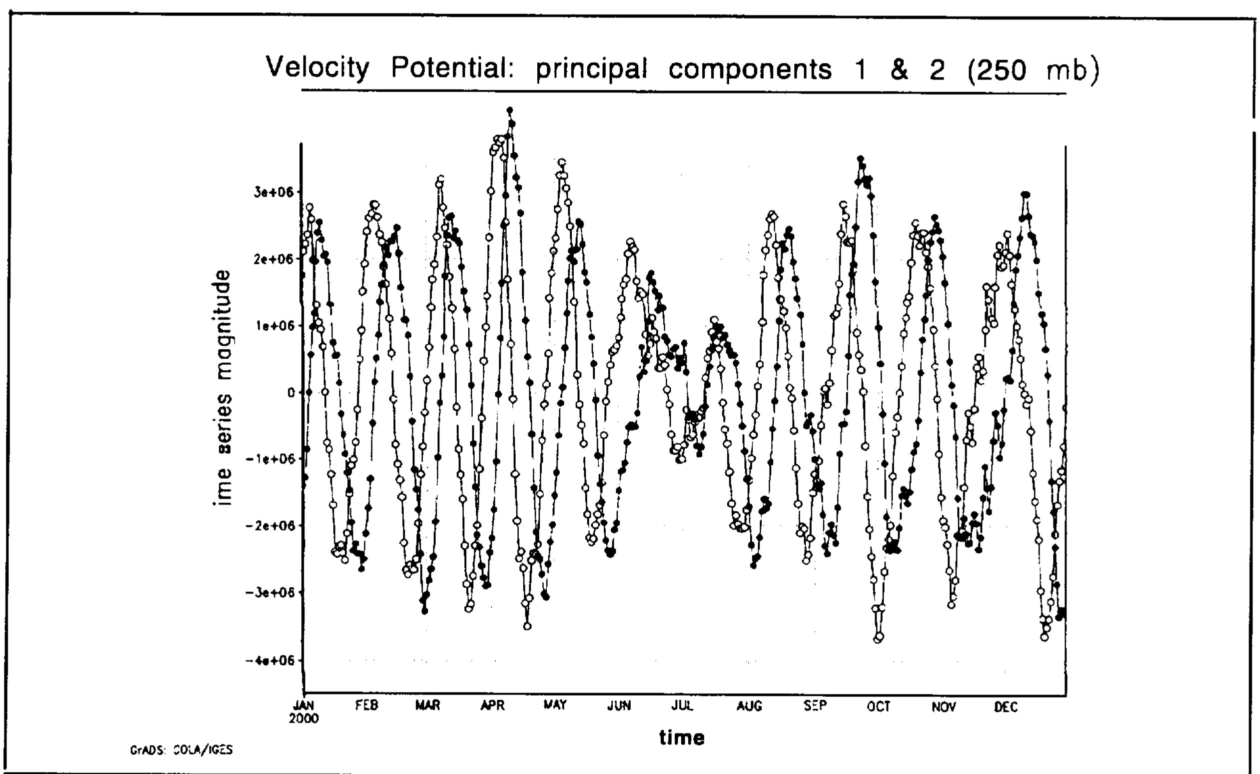


Figure 4. First (hollow circle) and second (solid circle) principal components of the MJO derived from χ_{250}.

4. Conclusion

The fact that the case that yielded the strongest MJO signal possessed the smallest temperature gradient between the pole and equator as well as the lowest equatorial temperatures raised the question as to which of these effects were responsible for the enhanced oscillation. A final experiment was performed in which the weakest gradient profile was shifted upward until temperatures at the equator corresponded to those of the original GISST distribution. It was found that the simulated fields matched those produced by the original GISST SSTs pointing to the possibility that for stable conditions, only damped moist Kelvin waves are generated. These waves are quickly dissipated for lack of a driving mechanism. Because the diabatic heating effectively reduces the static stability, and hence the rate of temperature change associated with vertical motions, moist Kelvin waves propagate at a reduced phase speed compared to dry waves-the stronger the heating, the slower the propagation (Bladé and Hartmann, 1993). This scenario would explain the weakened signal and slower phase speeds attained by SST profiles possessing warmer equatorial temperatures. Future research will examine transitional profiles for which dry Kelvin waves will be generated as well as looking exculsively at meridional wind velocity in order to gain a better understanding of the Rossby wave component associated with the oscillation. Additionally, longer three year runs will be performed in order to compile a larger statistical sampling of the MJO generated signals.

Acknowledgements. We would like to thank Howard University's Center for the study of Terrestrial and Extraterrestrial Atmospheres and the NASA Goddard Space Flight Center Laboratory for Atmospheres for support and technical assistance.

REFERENCES

Bladé, I. and D. L. Hartmann, 1993: Tropical intraseasonal oscillations in a simple nonlinear model. *J. Atmos. Sci., 50, 2922-2939.*

Chao, W. C. and Lin, S-J, 1994: Tropical intraseasonal oscillations, super cloud clusters, and cumulus convection schemes. *J. Atmos. Sci.*, **51, 1282-1297.**

, **1995:** A critique of wave-CISK as an explanation for the 40-50 day tropical intraseasonal oscillation. J. *Meteor. Sot. Japan, 48,677-683.*

Charney, J. G. and A. Eliassen, 1964: On growth of the hurrican depression. *J. Atmos. Sci.*, **21,68-75.**

Emmanuel, K. A., 1987: An air-sea interaction model of intraseasonal oscillations in the tropics. *J. Atmos. Sci.*, 44,2324-2340.

Gill, A. E., 1980: Some simple solutions for heat induced tropical circulation. *Quart. J. Roy. Meteor. Sot., 106, 447-463.*

Hyashi, Y.-Y., and A. Sumi, 1986: The 30-40 day oscillations simultated in an "aqua planet" model. *J. Meteor. Sot. Japan, 64,451-466.*

Held, I. M. and M. J. Suarez, 1978: A two level primitive equation model designed for climatic sensitivity experiments. *J. Atmos. Sci.*, 35,206-229.

Hendon, H. H., 1988: A simple model of the 40-50 day oscillation. J. *Atmos. Sci.*, 45,569-584.

Holton, J. R , 1992: An introduction to dynamic meteorology. Academic Press, 392 pp.

Krishnamurti, T. N. and D. Subrahmanyam, 1982: The 30-50 day mode at 850 mb during MONEX. J. *Atmos. Sci., 39, 2088-2095.*

Lau, K. M. and L. Peng, 1987: Origin of low-frequency (intraseasonal) oscillations in the tropical atmosphere. Part I basic theory. J. *Atmos. Sci.*, 47,950-972.

Lindzen, R. S., 1974: Wave-CISK in the tropics. J. *Atmos. Sci.*, **31,** 156-179.

Madden, R A. and P. R. Julian, 1971: Detection of a 40-50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, 28,702-708.

——— , and——— , 1972: Description of global-scale circulation cells in the tropics with a 40-50 day period. *J. Atmos. Sci.*, 29, 1109-1123.

— , 1986: Seasonal variations of the 40-50 day oscillation in the tropics. *J. Atmos. Sci., 43,3138-3158.*

Murakami, T., and T. Nakazawa, 1985, Tropical 40-50 day oscillations during the 1979 northern hemisphere summer. *J. Atmos. Sci., 42,* 1107-1122.

Neelin, J. D. , I. M. Held, and K. H. Cook, 1987: Evaporation-wind feedback and low frequency variability in the tropical atmosphere, *J. Atmos. Sci.*, 44, 2341-2348.

Rayner, N. A. et al, 1995: A New Global Sea-Ice and Sea Surface Temperature (GISST) Data Set for 1903-1994 for forcing Climate Models, Wadati Conference on Global Climate Change and the Polar Climate, Tsukuba, Japan.

Weickmann, K. M., 1983: Intraseasonal circulation and outgoing longwave radiation modes during Northern Hemisphere winter. Mon. Wea Rev., **111, 1838-1858,**

— , G. R. Lussky and J. E. Kutzbach, 1985: Intraseasonal (30-60 day) fluctuations of outgoing longwave radiation and 250 mb streamfunction during northern winter, Mon. Wea Rev., 113, 941-961.

URC97032 -32/51

The use of a chlorophyll meter (SPAD-502) for field determinations of red mangrove (*Rhizophora mangle* L.) leaf chlorophyll amount.

Xana M. Connelly
Department of Marine Sciences University of Puerto Rico

The red mangrove *Rhizophora mangle L.*, is a halophytic woody spermatophyte common to the land-sea interface of tropical and subtropical intertidal zones. It has been reported that 60 to 75% of the coastline of the earth's tropical regions are lined with mangroves (Walsh (1974) in Nybakken "1988). Mangroves help prevent shoreline erosion, provide breeding, nesting and feeding areas for many marine animals and birds. Mangroves are important contributors of primary production in the coastal environment, and this is largely proportional to the standing crop of leaf chlorophylls (Oswin and Kathiresan, 1994). Higher intensities of ultraviolet radiation, resulting from stratospheric ozone depletion, can lead to a reduction of chlorophyll in terrestrial plants (Tevini et al., 1980). Since the most common method for determining chlorophyll concentration is by extraction and this is labor intensive and time consuming, few studies on photosynthetic pigments of mangroves have been reported.

Chlorophyll meter readings have been related to leaf chlorophyll content in apples (Campbell et al., 1990) and maples (Sibley et al., 1996). It has also been correlated to nitrogen status in corn (Wood et al., 1992a) and cotton (Wood et al., 1992b). Peterson et al., (1993) used a chlorophyll meter to detect nitrogen deficiency in crops and in determining the need for additional nitrogen fertilizer. Efforts to correlate chlorophyll meter measurements to chlorophyll content of mangroves have not been reported. This paper describes the use of a hand-held chlorophyll meter (Minolta SPAD-502) to determine the amount of red mangrove foliar chlorophyll present in the field.

The chlorophyll absorption curve has peaks in the red (600-700 nm) and blue (400-500 nm) regions of the spectrum with little absorbance in the near-infrared region. The SPAD-502 measures transmittances in the red (650nm) where absorbance is high and unaffected by carotene and in the infrared (940 nm) region where absorbance by pigments is very low. Two light emitting diodes (LED's) (a red LED and an infrared LED) are built into the measuring head and emit light in sequence when the measuring head is closed. Light from these LED's goes through the emitting window, passes through the sample leaf in the measuring head, and enters the receiving window where it is converted into analog electrical signals. Then the ratio of the intensities of the transmitted light is calculated into a numerical SPAD value that is proportional to the amount of chlorophyll present in the leaf (Minolta Chlorophyll meter SPAD-502 Instruction Manual).

Healthy and mature leaves were obtained from mangroves growing at various localities along the southwest coast of Puerto Rico. SPAD measurements were taken by simply inserting a leaf and closing the measuring head, Six SPAD measurements were taken for each leaf and the average was calculated. Fresh leaf material was sampled using a cork borer (5mm diameter) and extracted immediately in pure acetone using an all-glass hand tissue grinder, following the procedure detailed by Corredor et al., (1995). Extracts were cleared by filtration (Gelman PTFE ACRODISC CR 0.2μm) and stored at -5°C until analysis. Absorbance of acetone extracts were measured at 644 and 662nm in a Hewlett-Packard 8452 spectrophotometer. The amount of chlorophyll was determined using the equations described by Lichtenthaler and Wellburn (1983). Regression analysis was used to assess the relationship between extracted and predicted chlorophyll values.

An absorption spectrum for *R. mangle* pigments extracted in pure acetone is shown in Figure 1, The maximum peaks characteristic of chlorophyll were observed at 430 and 662 nm.

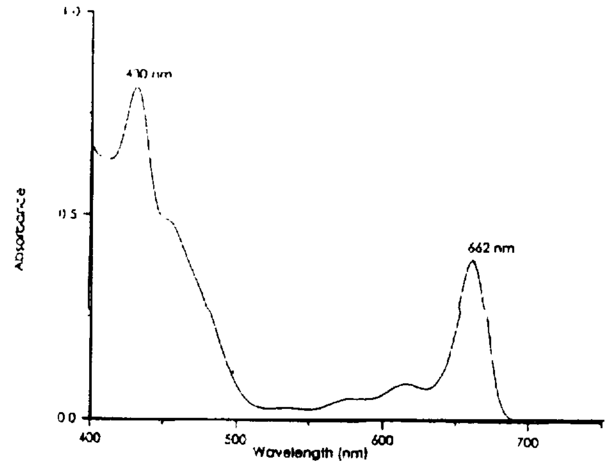


Figure 1. Red mangrove chlorophyll absorption spectrum in pure acetone extract

The chlorophyll meter, however, measures *in vivo* chlorophyll absorption which occurs at a slightly lower wavelength (650 nm) than ***in vitro*** extractions. The measured chlorophyll concentrations varied from 23.85 $\mu g/cm^2$ to 72.09 $\mu g/cm^2$, with an average of 43.74 $\mu g/cm^2$. The SPAD values ranged from 32.4 to 68.8. There is a high relationship between chlorophyll meter readings and tots\ chlorophyll (r^2 = 0.68) (Figure 2). The equation used to predict chlorophyll concentration from SPAD-502 readings is:

$$\text{Chlorophyll in } \mu g/cm^2 = (\text{SPAD} - 22.70) / 0.57$$

To validate the regression curve 14 samples were selected at random from the total of 52 samples. The relationship between the chlorophyll values obtained by extraction in pure acetone and the values predicted from SPAD-502 measurements is presented in Figure 3, Chlorophyll meter readings had a higher correlation with chlorophyll a (r* =0.77) than chlorophyll b (r^2= O. 56).

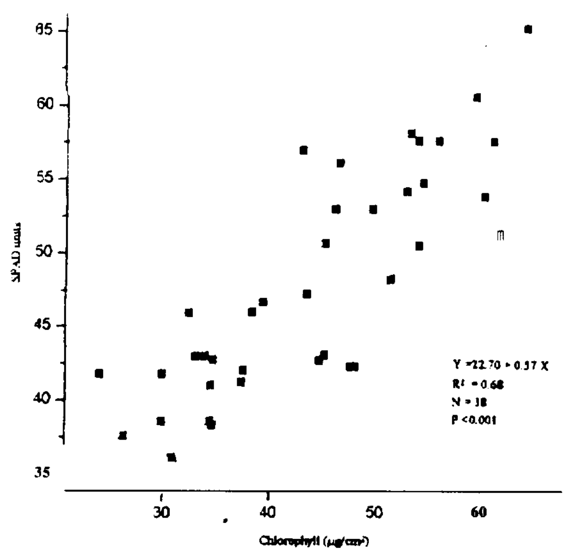


Figure 2. Relationship of Chlorophyll meter readings to chlorophyll extractions

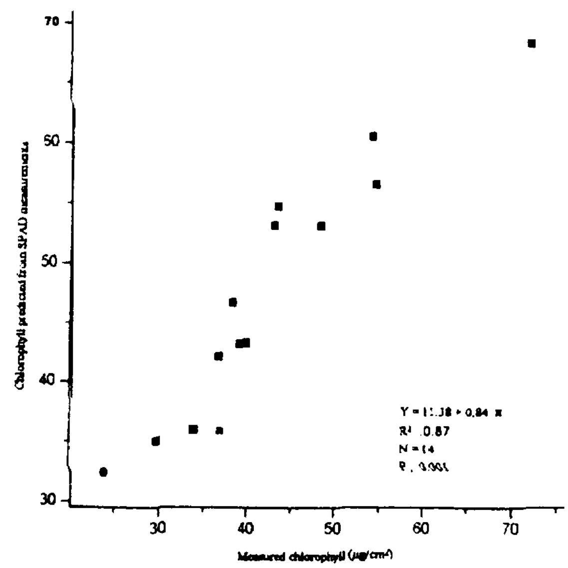


Figure 3. Relationship between predicted chlorophyll and extracted chlorophyll

Red mangrove leaves were collected from four stations in southwestern Puerto Rico in order to provide a wide range of chlorophyll values for the calibration. There were statistically significant differences between stations ($p<0.001$). These differences can be related to variations in the nutrient input received at each station. Means were significantly higher for the Sewage Treatment Plant (STP) station, with a mean chlorophyll value of 55.07 $\mu g/cm^2$. Intermediate values were found in Bird Island (42.39 $\mu g/cm^2$) and Caballo Blanco (39.77 $\mu g/cm^2$). Cayo Enrique had the lowest values with a mean of 37.34 $\mu g/cm^2$. Anthropogenic effects are responsible for the higher nutrient levels at the STP station, Bird Island has a large population of

birds which contribute high level of nutrients while Caballo Blanco and Cayo Enrique receive the least amount of exogenous nutrient imputs.

Another source of variation in the measured leaf chlorophyll values had to do with their exposure to solar radiation. Although the chlorophyll amount is expected to be reduced in sun leaves due to increased UV-B exposure (Tevini et al., 1981), there was no statistically significant differences ($p=0.05$) between the chlorophyll content of shaded leaves and sun exposed leaves, except for Caballo Blanco station. At this station, the chlorophyll content for sun exposed leaves had a significantly higher mean (44.57$\mu g/cm^2$) than the chlorophyll content for shaded leaves (mean of 34.29 $\mu g/cm^2$). At the other stations sun leaf chlorophyll values ranged from 26.47 $\mu g/cm^2$ to 72.09 $\mu g/cm^2$ while shade leaf values ranged from 23.85 $\mu g/cm^2$ to 62.04 $\mu g/cm^2$.

The chlorophyll a to b ratio varied between 3.19 in shaded leaves to 5.62 in sun exposed leaves, which agrees with Tevini et al., (1981) who reported that the proportion of chlorophyll a to chlorophyll b rises in UV-irradiated higher plants. However Kathiresan and Moorthy (1993) found that the chlorophyll a to b ratio was higher in shaded leaves ***of R. mucronata*** in which they report a chlorophyll a to b ratio of 1.3 for sun exposed leaves and 1.80 for shaded leaves.

In summary, the Minolta SPAD-502 Chlorophyll Meter was useful in providing a large number of leaf chlorophyll measurements in the field. In red mangroves, chlorophyll meter readings compared favorably to extracted values in pure acetone. In remote areas where traditional chlorophyll extraction procedures are impractical and samples can degrade before they can reach laboratory facilities, this portable meter can provide reliable measurements of the chlorophyll status of mangrove forests.

Acknowledgments

This work is a result of research sponsored by the National Aeronautic and Space Administration under grant NCCW-0088. I wish to acknowledge my major professor, Dr. Roy Armstrong for his assistance on the preparation of this paper, and J. Corredor and J. Morell for their helpful comments and laboratory facilities,

References

Campbell, R, J., K.N. Mobley, R.P. Marini and D.G. Pfeiffer, 1990. Growing conditions alter the relationship between SPAD-501 values and apple leaf chlorophyll. Hort Science 25 (3): 330-331.

Corredor, J. E., J.M. Morell, E.J. Klekowski and R. Lowenfeld, 1995. Mangrove genetics. HI. Pigment Fingerprints of chlorophyll-deficient mutants. Int. J. Plant Sci. 156 (1): 55-60.

Kathiresan, K. and P. Moorthy, 1993. Influence of different irradiance on growth and photosynthetic characteristics in seedlings of Rhizophora species. Photosynthetic 29(1): 143-146.

Lichtenthaler, H. K. and A. R. Wellburn, 1983, Determinations of total carotenoids and chlorophylls a and b of leaf extracts in different solvents. Biochemical Society Transactions 603: 591-592.

Minolta Chlorophyll meter SPAD-502 Instruction Manual 23pp.

Nybakken, J. W., 1988. Marine Biology: an ecological approach. 2nd edition, Harper& Row, Publishers, New York. pp **415-428.**

Oswin, S.D. and K. Kathiresan, 1994. Pigments in mangrove species of Pichavaran. Indian Journal of Marine Sciences. Vol. 23: 64-66.

Peterson, TA., TM. Blackner, DD. Francis and J.S. Schepers, 1993. Using a chlorophyll meter to improve N management. NebGuide, Cooperative Extension, Institute of Agriculture and Natural Resources, University of Nebraska 4 pp.

Sibley, JL., DJ. Eakes, CH. Gilliam, GJ. Keever, W.A. Dozier and DG. Himelrick, 1996. Foliar SPAD-502 meter values, nitrogen levels, and extractable chlorophyll for red maple selections. HortScience 31 (3): 468-470.

Tevini, M., W. Iwanzik and U. Thoma, 1980. The effects of UV-B radiation ***on*** higher plants. in Calkins, J. (Ed.), 1980. The role of ultraviolet radiation in marine ecosystems. Plenum Press, New York and London. pp 581-6"16.

Tevini, M., W. Iwanzik and U. Thoma, 1981. Some effects of enhanced UV-B radiation on the growth and composition of plants. Planta 153: 388-394.

Wood, C. W., D. W. Reeves, R.R. Duffield and K.L. Edmisten, 1992a. Field chlorophyll measurements for evaluation of corn nitrogen status. Journal of Plant Nutrition 15 (4): 487-500.

Wood, C. W., P. W. Tracy, D.W. Reeves and K.L. Edmisten, 1992b. Determination of cotton nitrogen status with a hand-held chlorophyll meter. Journal of Plant Nutrition 15 ***(9): 1435-1448.***

GROWTH, SPECTROSCOPY AND PHOTOREFRACTIVE INVESTIGATION OF VANADIUM DOPED CdSSe

Michelle Davis, Zhengda Pan, Kuo-Tong Chen, Henry Chen, Swanson L. Davis, Arnold Burger, and Steven H. Morgan,

Center for Photonic Materials and Devices, Physics Department, Fisk University, Nashville, TN 37208

We present two-wave mixing results obtained with a $CdS_{0.8}Se_{0.2}$:V crystal. The $CdS_{0.8}Se_{0.2}$:V crystal was grown by physical vapor transport (PVT) along with a concentration of 150 ppm (nominal) vanadium for creating trap centers. The as-grown crystal has a large crystal size, good optical quality, and a medium resistivity of 10^6 -108 Ω-cm. A large photorefractive gain coefficient of ().24 cm-' was observed at 633 nm with an optical intensity of 60 mW/cm^2 and a grating period of 1.6 μm. To our knowledge, this is the first observation of the photorefractive effect in a vanadium doped CdSSe crystal. Room temperature absorption and low temperature photoluminescence spectroscopy measurements are also discussed. With a significant photorefractive effect, the CdSSe:V crystals are promising for device applications based on photorefractive effect, in the wavelength range of 600-700 nm.

Introduction

Photorefractive materials are nonlinear optical materials which experience a significant change in the refractive index when expo sed to inhomo geneo us illumination. These materials have potential device applications in optical signal processing and related areas, including reversible holographic storage, tracking filters, optical interconnects, etc. Photorefractive crystals are being used to record the holographic interconnections between the neurons in optical neural networks. The optical implementation of artificial neural networks are useful in solving problems such as pattern recognition and robotics. Semiconductor crystals such as CdTe:V, GaAs, InP:Fe were reported as having fast response times at low incident intensities of few tens of mW/cm^2. The reported value of the electro-optic coefficient of CdTe:V is three times larger than that of GaAs and InP. However, the above materials can not be used in the wavelength range of 600 to 700 nm. [1-3]

The as-grown $CdS_{0.8}Se_{0.2}$ crystal has a large crystal size, good optical quality, medium to high resistivity, and a good transparency for wavelengths longer than 600" nm. The ternary system offers the capability that its band-gap can be tailored by adjusting the alloy composition, thus the crystals may be optimized for use at desired wavelengths.

It has been reported that the vanadium dopant is responsible for the high photorefractive gain observed in CdTe:V. Vanadium dopant in CdTe compensates for a high resistivit y, and provides additional deep levels. [1] We assume that vanadium dopant in CdSSe would have behave similarly to the vanadium dopant in CdTe.

We present here the first results of photorefractive two-beam coupling measurements on a $CdS_{0.8}Se_{0.2}$ single crystal doped with vanadium. The measurements were performed at 633 nm using a 5 mW cw He-Ne laser. We also present the photoluminescence and absorption spectra. These results are discussed and compared to that of vanadium-doped CdTe crystals and a pure CdS crystal reported previously [1-4].

Theory

In a photorefractive two-beam coupling experiment (Fig. 1), two beams of unequal intensity intersect inside the semi-insulating sample to forma spatially sinusoidal interference pattern. in the region of constructive interference the charge carriers are excited to the band states, undergo diffusion, drift, and are recaptured by traps. These effects result in charge redistribution and give rise to the space-charge field. This field, acting through the linear electro-optic effect, modulates the refractive index of the material. If the grating is displaced with respect to the incident optical interference pattern the beam coupling gives rise to energy transfer from one beam to the other. The grating wave-vector depends on the wave-vectors of the incident beams and the angle between the two crossed beams. The value of grating wave-vector is given by

$$k_g = 2k\sin\theta \tag{1}$$

where 2θ is the angle between the two crossed beams.

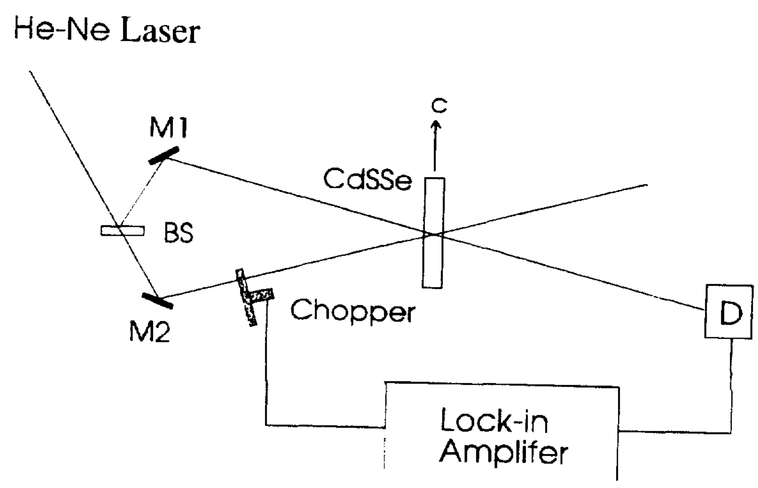


Figure 1: Experimental setup for photorefractive two-beam coupling.

By chopping the pump beam and monitoring the signal beam, the photorefractive gain can be measured. The ratio of the modulation of the signal-beam ΔI_s to the signal beam I_s without pump beam is related to the photorefractive gain Γ. The relation is given by [4]

$$\frac{\Delta I_s}{I_s} = \frac{\exp(\Gamma L) - 1}{1 + \beta \exp(\Gamma L)} \qquad (2)$$

where L is the overlap length between two beams and β is the intensity ratio of the signal beam to pump beam.

According to a simplified model given by Kukhtarev and others, the photorefractive gain Γ can be predicted by the formula [1]

$$\Gamma = \frac{2\pi (n_0^3 \gamma_{eff})}{\lambda \cos\theta} \left(\frac{KT}{e}\right) \frac{k_g}{1 + (k_g^2/k_D^2)} \xi \qquad (3)$$

where ξ is the electron-hole competition factor, k_D is the inverse of the Debye screening length, and γ_{eff} the effective electro-optic coefficient, For an one level model, $k_D^2 = (e^2/\epsilon KT)\, N_{eff}$ where N_{eff} is the effective trap density and ϵ is the dielectric constant. The γ_{eff} depends not only on electro-optic properties of the crystal, but also on the sample orientation and beam polarizations, this relation is

$$\gamma_{eff} = \hat{e}_1^*(Rk_g)\, \hat{e}_2 \qquad (4)$$

where R $(3\times3\times3)$ is the linear electro-optic tensor, k_g (3×1) is the unit grating wave-vector, $\hat{e}_1{}^*$ (1×3) and $\hat{e}_2$ (3×1) are the unit polarization vectors of signal and pump beams, respectively. CdSSe belongs to the hexagonal 6 mm symmetry group with nonzero electro-optic coefficients, $\gamma_{13} = \gamma_{23}$, γ_{33} $(\gamma_{42}) = \gamma_{51}$. For the experimental configuration used in this study, $\gamma_{eff} = \gamma_{13}$.

Results and Discussion

A single crystal of $CdS_{0.8}Se_{0.2}$:V was grown by horizontal physical vapor transport (PVT) method along with a concentration of 150 ppm (nominal) vanadium for creating trap center. The starting materials were supplied by Cleveland Crystals, Inc., with a labeled purit y of 99 .995°/0. The as-grown crystal had a diameter of 12 mm and a length of 6 cm. A sample from this crystal was cut and polished to a $8\times5\times4$ mm^3 rectangular parallelepipeds with the orientation shown in Fig. 1. The resistivity was measured by applying silver-paint contacts on the 3 x 3 mm^2 area and measuring the I - V curve. The resistivity is in excess of 10^7 Q-cm at voltages up to 100 V.

The absorption spectra measured at room temperature is shown in Fig. 2. The spectrum shows a absorption edge at 585 nm and a long tail. The absorption edge corresponds to the bandgap of the semiconductor crystal and the absorption tail is assumed to be from the vanadium dopant. An absorption coefficient of 2.3 cm-1 was obtained at 633 nm. This value is comparable to that of a CdTe:V sample with vanadium concentration 5 $\times 10^{16}$ $/cm^3$ [3].

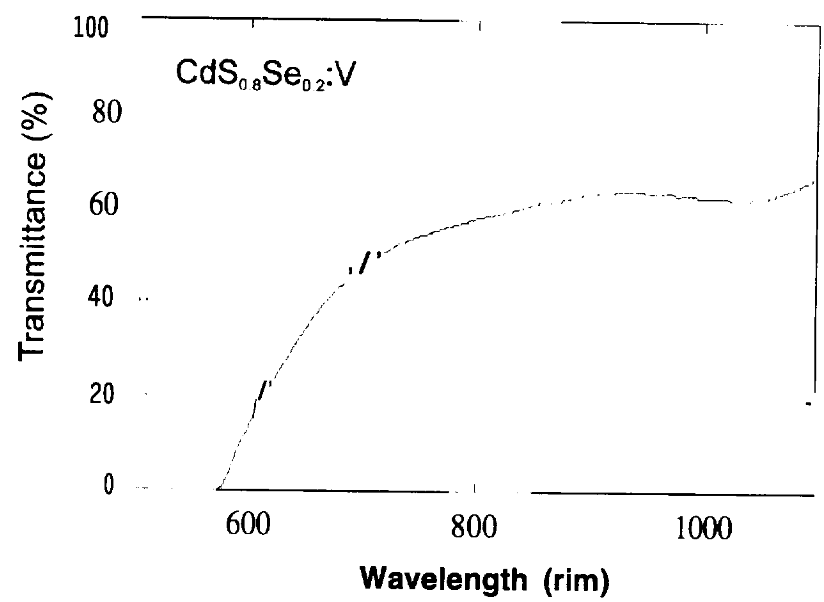


Figure 2: Room temperature absorption spectrum of $CdS_{0.8}Se_{0.2}$:V.

Photoluminescence (PL) measurements were performed at 5.8 K. Fig. 3 shows the PL spectrum of a $CdS_{0.8}$ $Se_{0.2}$:V crystal. A strong bound exciton peak is located at 2.313 eV with two phonon replicas at 2.279 and 2.244 eV. The spacing of the phonon replicas is about 35 meV, which is between the longitudinal (LO) phonon of 38 meV in CdS and the LO phonon of 24 meV in CdSe [6]. The broad band from 1.7 to 2.1 eV (centered at 1.925 eV) is attributed to a defect center. Similar emission band was observed in a vanadium-doped CdTe, but the exact nature of the centers

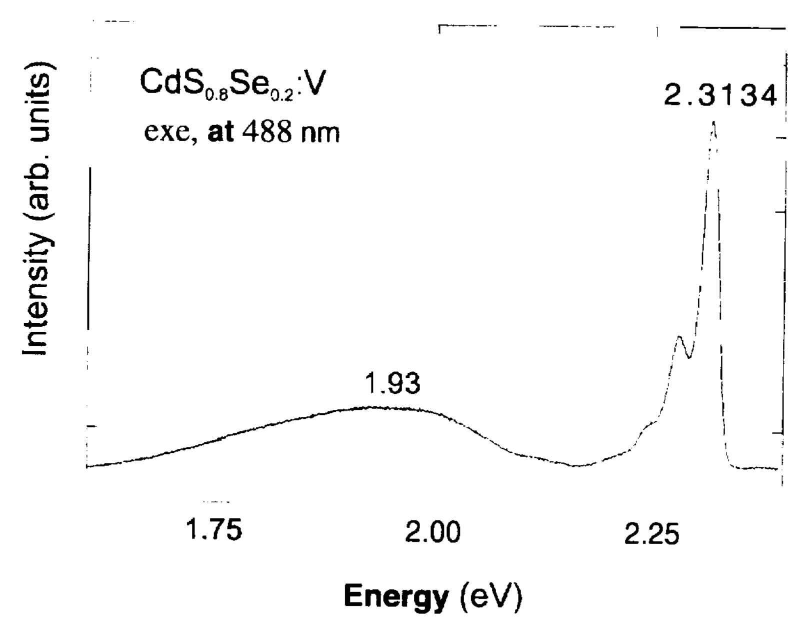


Figure 3: Photoluminescence spectrum at T = 5.8K.

responsible for these emissions is still unknown or is a matter of speculation [5]. It should be noticed that the wavelength of 633 nm (1 .959 eV) from the He-Ne laser matches well with this impurity band observed in PL spectrum.

The photorefractive properties were studied by two-beam coupling at 633 nm using a 5 mW cw He-Ne laser. Two s-polarized Gaussian beams (beam size is about 1.2 mm in diameter) with a beam intensity ratio of $\beta = I_s(z=0)/I_p(z=0) = ().41$ interfered inside the crystal. The grating wave vector was oriented along the c-axis of the crystal. The pump beam was modulated at84 Hz by a chopper. The beam-coupling gain Γ reached a value of about 0.24 cm^{-1} at a maximum intensity of 60 mW/cm^2 (I_p) and a grating spacing around 1.6 pm. This gain coefficient is similar to the value reported for a CdTe:V sample at 1.32 µm with an intensity of 10 mW/cm^2 [3] and for a pure CdS crystal at 633 nm with an intensity of 80 mW/cm^2 [4]. The value of photorefractive gain decreases significantly with decreasing laser intensity. The Γ value decreased to one fourth for an incident intensity of 30 mW/cm^2. Our $CdS_{0.8}Se_{0.2}$:V sample has a relatively lower resistivity (1 0^7 Ω-cm) than that of a pure CdS crystal (10'0 Ω-cm), a higher photorefractive gain therefore could be expected at a higher laser intensity when the photo excited carrier concentration exceeds the thermally excited carrier concentration. The gain coefficient exhibits a large fluctuation at different points on the sample. This implies some degree of inhomogeneity in our sample. We also performed the two-beam coupling experiment at 800 nm by using a cw Ti:sapphire laser, but no gain signal beyond noise can be obtained. The results we observed are therefore attributed to the resonant enhancement of the photorefractive response due to photo-excitation near the band edge of the crystal.

Conclusion

We have reported the first observation of the photorefractive effect in a $CdS_{0.8}Se_{0.2}$:V crystal at wavelength 633 nm using a cw He-Ne laser. The observed photorefractive gain is comparable to that of the best photorefractive semiconductors. The $CdS_x Se_{-1-x}$:V crystals maybe fabricated with a large size, a high resistivity, and controllable defect densities. This material therefore has potential as a fast and sensitive photorefractive material in the wavelength range of 600 to 700nm.

Acknowledgments

This Research is supported by the NSF Grant HRD-9550605, US Air Force Grant GFY1457096-03120, and NASA Grant NAGW-2925.

References

[1] R. B. Bylsma, P. M. Bridenbaugh, D. H. Olson, and A. M. Glass, "Photorefractive properties of doped cadmium telluride", Appl. Phys. Lett. 51, 889 (1 987).

[2] A. Partovi, J. Millerd, E. M. Garmire, M. Ziari, W. H. Steier, S. B. Trivedi, and M. B. Klein, "Photorefractivity at 1.5µm in CdTe:V", Appl. Phys. Lett. 57, 846(1990).

[3] J.-Y. Moisan, N. Wolffer, O. Moine, P. Gravey, G. Martel, A. Aoudia, E. Repka, Y. Marfaing, and R. Triboulet, "Characterization of photorefractive CdTe:V: high two-wave mixing gain with an optimum low-frequency periodic external electric field', J. Opt. Sot. Am. B 11, 1655(1994).

[4] P. Tayebati, J. Kumar, and S. Scott, "Photorefractive effect at 633 nm in semi-insulating cadium sulfide", Appl. Phys. Lett. 59, 3366 (1991).

[5] R. N, Schwartz, M Ziari, and S. Trivedi, "Electron paramagnetic resonance and an optical investigation of photorefractive vanadium-doped CdTe", Phys. Rev. B 49, 5274 (1994).

[6] A. K. Arora and A. K. Ramdas, "Resonance Raman scattering from defects in CdS", Phys. Rev. B, 4345 (1987).

URC97034

An Approach to Building a Traceability Tool for Software Development

Nelly Delgado and Tom Watson
Department of Computer Science
The University of Texas at El Paso
El Paso, TX 79968
ndelgado, twatson@cs.utep.edu

1 Introduction

It is difficult in a large, complex computer program to ensure that it meets the specified requirements. As the program evolves over time, all program constraints originally elicited during the requirements phase must be maintained. in addition, during the life cycle of the program, requirements typically change and the program must consistently reflect those changes. Imagine the following scenario. Company X wants to develop a system to automate its assembly line. With such a large system, there are many different stakeholders, e.g., managers, experts such as industrial and mechanical engineers, and end-users. Requirements would be elicited from all of the stakeholders involved in the system with each stakeholder contributing their point of view to the requirements. For example, some of the requirements provided by an industrial engineer may concern the movement of parts through the assembly line. A point of view provided by the electrical engineer may be reflected in constraints concerning maximum power usage. End-users may be concerned with comfort and safety issues, whereas managers are concerned with the efficiency of the operation. With so many points of view affecting the requirements, it, is difficult to manage them, communicate information to relevant stakeholders, and it is likely that conflicts in the requirements will arise. In the coding process, the implementors will make additional assumptions and interpretations on the design and the requirements of tile system. During any stage of development, stakeholders may request that a requirement be added or changed. in such a dynamic environment, it is difficult to guarantee that the system will preserve the current set, of requirements.

Tracing, the mapping between objects in the artifacts of the system being developed, addresses this issue. Artifacts encompass documents such as the system definition, interview transcripts, memoranda, the software requirements specification, user's manuals, the functional specifications, design reports, and system code. Tracing helps 1) validate system features against, the requirement specification, 2) identify error sources and, most importantly, 3) manage change [4]. With so many people involved in the development of the system, it becomes necessary to identify the reasons behind the design requirements or the implementation decisions.

This paper is concerned with an approach that maps documents to constraints that capture properties of and relationships between the objects being modeled by the program. Section 2 provides the reader with a background on traceability tools. Section 3 gives a brief description of the context monitoring system on which the approach suggested in this paper is based. Section 4 presents an overview of our approach to providing traceability. The last section presents our future direction of research.

2 Background

The typical approach to maintaining traceability, especially for complex systems, requires that all system artifacts created at various stages of the development process be linked to the requirements [5]. In such an approach, there must be hyperlinks (physical links) between all artifacts and requirements. These links should provide bidirectional, vertical and horizontal traceability. Bidirectional *traceability* refers to the ability to trace both backward and forward. *Vertical traceability* allows the user to trace between documents developed from different life cycles, whereas, *horizontal traceability* refers to the links between related objects created in the same life cycle [5]. Each document must have a logical structure so that the tracing tools will understand the interrelationships between different software documents and have the ability to update the links as the

system evolves [6]. With a large volume of documents, however, it is difficult to maintain and update the links between the artifacts.

Through the links, tracing can also provide information concerning accountability for requirements, design, and implementation decisions, The ability to track projects and manage design rationale are other uses for traceability. The rest of this section deals with the traditional approach to tracing, i.e., tracking the requirements to the implementation.

Some approaches to building a traceability tool include: an object-oriented approach, a graph-based approach, and an approach that involves the management through a project database. In the object-oriented approach, users define the classes of artifacts and the relationships between them. The classes of artifacts provide the logical structure necessary for the documents to be traced. The relationship classes provide the structure for defining links and their relations. The use of relations instead of simple links lets developers distinguish among different links between the same objects. Also, by using properties of relations, this approach can relate objects that are not directly linked [4].

The graph-based approach takes both the coarse-grain level and the fine-grain level of the system into account. On the coarse-grain level, links represent dependencies between whole documents. On the fine-grain level, the structure of the documents are taken into account. Both levels are necessary to provide for adequate traceability. The collection of documents is represented by the use of hierarchical graphs. Operations on the graphs are defined by means of a formal language based on a graph rewriting system [6], On the coarse-grain level, ail project documents and the relations between the documents are represented in the graph, On the fine-grain level, links represent the relation between individual objects in documents. These objects may be contained in the same document or in separate documents.

The project database model involves a database management system and an object-based model of software life cycles. All of the documents created during the development of the system are stored in a project database [3]. Similar to previous approaches, this approach requires that documents and relations are highly structured. Using a predefine document structure and a set of document relationships, documents are developed to allow the management of links. Through the use of a database, stored documents can be written in either natural or formal languages. Key words and key elements need to be identified as the user creates the documents to provide points for tracing across documents [3].

These approaches to developing traceability tools all have one thing in common - they link all artifacts to requirements. The advantage to having direct links' between all artifacts is that one can trace the system and ensure that the system requirements have been met, thus providing more complete coverage of the system. In a system with thousands of interrelated objects, however, getting all of the information available may not, be useful. Too much information could provide the user with an over abundance of irrelevant data [4].

3 An Overview of the Context Monitoring System

Context monitoring is an approach for managing properties and relationships on data of a program. Context monitoring consists of the following:

- the specification of integrity constraints on data or objects being modeled by a software system, and
- a constraint satisfiability mechanism that verifies their enforcement during the program's execution.

Integrity constraints are the conditions that data maintained in a knowledge base must satisfy as it evolves. Specified in a typed first-order logic language [2], they formalize properties about data objects and the program.

Fig. 1 provides an overview of the context, monitoring system. The constraints capture properties, relationships, restrictions and limitations on data or objects of the program and may be used to reflect the interpretations and assumptions that, are made about the objects during development. The constraints are elicited from numerous sources, denoted by S_i in the figure, that includes the customer, domain experts, analysts, and members of the design, implementation and maintenance team.

The constraints reside in a repository with links to the documentation. This may be in the form of a data dictionary, requirements definition, software requirements specification, design document, user's manual, program documentation, memoranda, interview transcripts, or videotape transcripts. The documents may reside in a project database or maintained as separate text files. If a document is not kept on-line, then the linked file would contain merely the location of the document and the location of the constraint in that document,

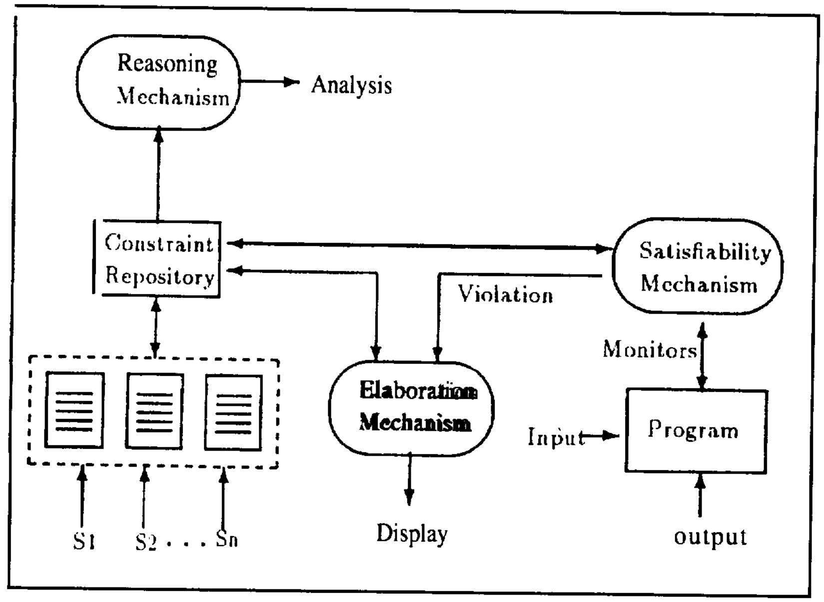


Figure 1: Overview of the context monitoring system.

The constraint satisfiability mechanism provides dynamic verification that, the actual behavior of the program corresponds to the intended behavior of the program. When a constraint violation occurs, the monitoring mechanism captures the state of the machine, both prior to the violation and at the time of the violation. A violation may be an indication of one of the following:

- The program does not maintain the constraint.
- The input data does not meet specified properties.
- Conflicting constraints exist, which cause at least one of the constraints to be violated.
- A change to the program has violated an existing constraint,
- The environmental context in which the program is running has changed, invalidating the assumption made while designing and coding the program.

The user is supplied with information concerning the Violation through the elaboration mechanism. Because the documents are linked to the constraints, the user has access to the source(s) of the constraint, justifications, and other information that, may be contained in the documentation. The reasoning mechanism is a tool that groups related constraints for analysis and determines potential inconsistencies between constraints apart from the program itself. This permits static analysis of the knowledge collected from multiple sources before too much time has been invested in the design and/or implementation.

It is important to note that with this approach, the constraints are not embedded in the program, but are maintained independent of the program. Separating the integrity constraints ensures that changes in the code do not inadvertently change a constraint and, through constraint satisfiability, that added or changed code does not violate existing constraints.

4 Traceability Using Context Monitoring

The approach suggested here provides a way for linking between requirements, constraints, and parts of the program that, manipulate constrained variables. A distinguishing characteristic is that the developer does not have to create physical links between requirements and the implementation. 'This simplifies tracing because constraints are not embedded in the program and, as a result, the user does not have to worry about changes in the code that may alter links. Tracing in the context monitoring system only involves physics!

links between the constraints stored in the repository and the documents as shown in Fig. 2. Although this approach does not provide links between all requirements and program segments, it does provide a way to link a significant subset of the requirements.

Another characteristic of this traceability approach addresses fluctuating and conflicting requirements, a major issue in software development [I]. Consider a case where conflicting constraints are specified on a requirement. In such a case, the satisfiability mechanism will detect a violation if it is impossible for the program to satisfy both requirements. Through traceability, the user will have access to the sources of the constraint providing a basis for resolving the conflict. In addition, the traceability tool is not, preoccupied with managing constraints. Because the constraints are not embedded in the program, even if implementation completely changes, the links between constraints and the documents remain the same unless the constraints are deleted.

The approach presented in this paper provides the means to physically link the constraints to the various requirement documents through bidirectional hyperlinks. Because only those parts of the requirements and documents that specify a constraint are linked to the corresponding formal constraint and links are not maintained in the code, the number of links to be managed is reduced. In other words, constraints should always be tied directly to a document.

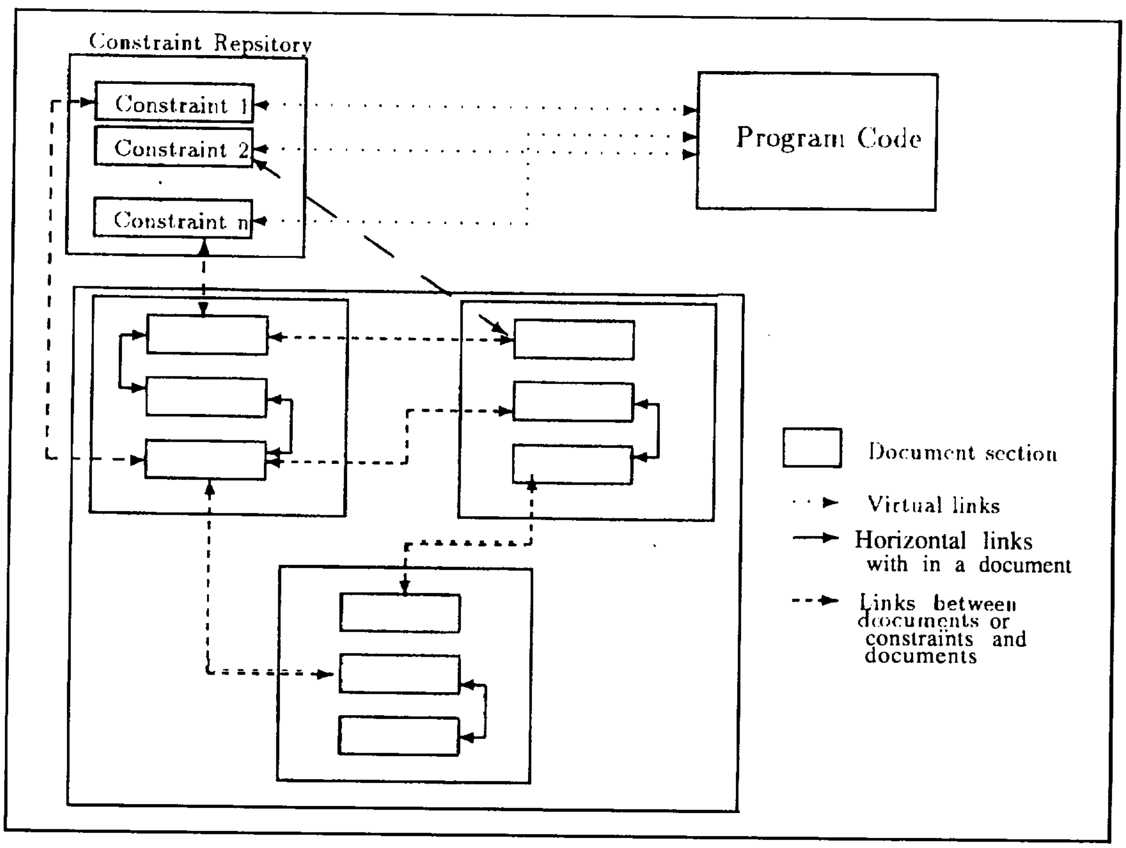


Figure 2: Tracing in the context monitoring environment.

The context monitoring system provides for bidirectional, vertical and horizontal traceability, and also allows the user to check for accountability, design rationale and dependencies. Project tracking, however, cannot be done with this approach. Two situations may occur. One is that a violation occurs during program execution and the user wants to determine the source of the constraint. Through the elaboration mechanism, it is possible to return to the repository based on the constraint, violated; therefore, the user can link back to the document. A second situation is when the user requests information concerning the impact of a requirement, or a constraint on a program. In this case, the satisfiability mechanism gives information on the code that is affected by this constraint, providing a virtual link between the constraint. and the code. Similarly, this procedure can be followed to go from the code to the constraint repository. Clearly, the links

between the constraints and the documents are fine-grain and the virtual links between the constraints and the program code are coarse-grain, (making the links between the code and any documents coarse-grain as well). As a result, any queries on the coarse-grain links will lead the user to candidate areas of interest in either the code or the documents. Recall that the relationship between these documents is based on the constrained variables.

Related constraints from different documents are linked together providing both horizontal and vertical traceability. This allows the user to determine if a constraint is supported by more than one document or decision. These links may be contained in the same document or across multiple documents providing horizontal traceability with a fine-grain level.

Because all constraints in the system can be linked back to a decision in the documentation through vertical links, design rationale can be easily traced. Provided that the documents contain the information on whom originally specified the requirement, the resulting constraint can be linked to this requirement, providing accountability to the decision makers.

5 Summary

The traceability tool in the context monitoring system allows virtual links from the code to the constraints providing tracing between requirements that specify relationships between properties of objects on the system. The advantage to this approach is that the system does not, have to provide hard links directly to the code; therefore, when modifications are made in the code, links are not inadvertently affected. Additionally, this allows the user to trace only the relevant information corresponding to a constraint. One disadvantage to this approach is that it, is not possible to ensure that all system requirements are met because the constraints capture only those parts of the requirements that specify properties and relationships between objects.

Future research in the area of traceability tools in context monitoring includes many different goals. First, we must determine the best method for storing the documents of the system. As shown in the discussion of traceability tools in section 2, this must be a highly structured approach to allow the system to process the document links. Some possible approaches to providing this structure include the object-oriented, the graph-based and the project database management approach.

The next objective is to develop a prototype to model the proposed traceability tool. This prototype will provide the tracing technique to the context monitoring system. The elaboration mechanism, constraint repository and documents will be included in the implementation of this prototype.

Acknowledgments. This work was supported by NASA under contracts NAG-1012 and NCC W-0089, and NSF grant no. CDA-9522207.

References

[1] Curtis, B., H. Krasner and N. Iscoe, "A Field Study of the Software Design Process for Large Systems", *communications of the ACM, 31* (11), pp. 1268-1287. Nov. 1988.

[2] Gates, A. and F. Fernandez, "Building Systems with Integrity Constraints," to be published in *The Proceedings of the Second World Conference on Integrated Design and Process Technology*, Dec. 1-4, 1996, Austin, Texas.

[3] Horowitz, E. and R. Williamson, "SODOS: A Software Documentation Support Environment-Its Use", *IEEE Transactions on Software Engineering*, SE- **12**(11), pp. 1076-1087, Nov. 1986.

[4] Pinheiro, F. and J. Goguen, "An Object-Oriented Tool for Tracing Requirements", *IEEE Software*, pp. 52-64, Mar. 1996.

[5] Ramesh, B. and M. Edwards, "Issues in the Development of a Requirements Traceability Model", *Proceedings of the IEEE International Symposium on Requirements Engineering*, San Diego, CA: IEEE Computer Society Press, 1993, pp. 256-259.

[6] Westfechtel, B., "A Graph-Based Approach to the Construction of Tools for the Life Cycle", *IEEE*, pp. 2-13, June 1992.

Page intentionally left blank

S35/54

URC97035

Evaluation Of A Human Modelling Software Tool In The Prediction Of ExtraVehicular Activity Tasks For An International Space Station Assembly Mission.

H. Charles Dischinger, Jr. and Tomas E. Loughead
NASA Marshall Space Flight Center, Alabama 35812
c.dischinger@msfc.nasa.gov
Sigmatech, Inc., Huntsville, Alabama

KEYWORDS: Human modelling, extravehicular activity, International Space Station, neutral buoyancy

ABSTRACT

The difficulty of accomplishing work in extravehicular activity (EVA) is well documented. It arises as a result of motion constraints imposed by a pressurized spacesuit in a near-vacuum and of the frictionless environment induced in microgravity. The appropriate placement of foot restraints is crucial to ensuring that astronauts can remove and drive bolts, mate and demate connectors, and actuate levers. The location on structural members of the foot restraint sockets, to which the portable foot restraint is attached, must provide for an orientation of the restraint that affords the astronaut adequate visual and reach envelopes. Previously, the initial location of these sockets was dependent upon the experienced designer's ability to estimate placement. The design was tested in a simulated zero-gravity environment; spacesuited astronauts performed the tasks with mockups while submerged in water. Crew evaluation of the tasks based on these designs often indicated the bolt or other structure to which force needed to be applied was not within an acceptable work envelope, resulting in redesign. The development of improved methods for location of crew aids prior to testing would result in savings to the design effort for EVA hardware. Such an effort to streamline EVA design is especially relevant to International Space Station construction and maintenance. Assembly operations alone are expected to require in excess of four hundred hours of EVA. Thus, techniques which conserve design resources for assembly missions can have significant impact. We describe an effort to implement a human modelling application in the design effort for an International Space Station Assembly Mission. On Assembly Flight 6A, the Canadian-built Space Station Remote Manipulator System will be delivered to the U.S. Laboratory. It will be released from its launch restraints by astronauts in EVA. The design of the placement of foot restraint sockets was carried out using the human model Jack, and the modelling results were compared with actual underwater test results. The predicted locations of the sockets was found to be acceptable for 94% of the tasks attempted by the astronauts. This effort provides confidence in the capabilities of this package to accurately model tasks. It therefore increases assurance that the tool may be used early in the design process..

INTRODUCTION

The National Aeronautics and Space Administration has a long history of application of rigorous methodology to systems engineering. The approach extends back to the very beginning of the von Braun era and thus predates the wide adoption of design principles that have since been applied in, for example, computer systems (*cf.* Gould and Lewis, 1985). The principles can be summarized as: a) an early focus on the requirements of the system, b) testing to ensure the system meets the requirements, and c) iterative design. In the case of human-occupied and operated space hardware, the requirements include a focus on the needs of the user. The system must support or be operable by the astronaut. Otherwise, the other system components will be of little value, however well designed. One of von Braun's contributions to the empirical phase of design of human space systems is the introduction of underwater testing to NASA at the Marshall Space Flight Center. This test facility, known as the Neutral Buoyancy Simulator, and others like it are used to provide a frictionless environment in which engineers and astronauts can assess design effectiveness. Mockups of the hardware are submerged and outfitted with flotation to make them neutrally buoyant; *i.e.*, they neither sink nor float, but tend to retain their position, as an object would in space. The test subject, likewise deprived of friction, is able to evaluate whether useful work (e.g., bolt drive, connector mate and demate, and lever actuation) can be accomplished with the system. This type of testing is especially useful for ExtraVehicular Activity (EVA) hardware. The motion constraints imposed by a pressurized spacesuit in a near-vacuum are difficult to simulate outside of the underwater environment, due to the weight of the suit. These constraints, along with the effects of suit bulk, must be incorporated in the testing to ensure the design compensates for them.

Construction of the International Space Station is expected to require in excess of four hundred hours of EVA. Neutral buoyancy testing, while effective when properly integrated with the other design principles, is expensive. The cost of testing all of the designs which must support this EVA is a significant part of the design budget. Methods which allow the proper implementation of system design principles while reducing this cost would be of considerable value. In the last ten years, there has been a near-complete transition in the engineering design environment from two-dimensional drawings to three-dimensional Computer-Aided Design (CAD). The human factors personnel responsible for ensuring EVA designs are workable are thus afforded

new opportunities to examine the human/machine interfaces through computer simulation. Computer models allow more thorough insight into these interfaces than was previously possible with drawings. Several human modelling packages have been developed which allow simulation of the work tasks as they would be performed in space, the human figure is placed into position to interact with the CAD models of the hardware. However, validation of these modelling systems must be accomplished before their results can be accepted by designers. Only a few efforts have been made to compare the computer simulation results with real-life human/machine interaction. Since the simulation mimics the test, as well as the space environment, neutral buoyancy testing provides the opportunity to examine the results of computer modelling. This paper will examine the utility of a human modelling software application in the examination of tasks associated with an International Space Station EVA design. The work has two components: a modelling effort and a comparison between the predictions from modelling and the results of neutral buoyancy evaluation of the tasks.

The design effort for which the human factors engineer is primarily responsible in EVA-operated hardware is the reduction of the effects of microgravity. Useful work is essentially impossible for a free-floating astronaut. One solution to this problem developed by NASA is to secure the astronaut on a platform called a foot restraint. The astronaut slides his or her boots into loops on a foot restraint (see Figure 1), securing them to the plate of the restraint. A foot restraint thus provides anchorage against which to react the forces required to release or tighten bolts and perform other tasks. The appropriate placement of foot restraints is critical to accomplishment of EVA work tasks, and it is this placement which comprises a significant portion of the human factors engineer's contribution to the design. This placement of the restraints is also a useful metric for the validation of the human modelling packages. The bulk of the modelling effort in this first attempt to assess the efficacy of the software consisted of positioning the restraints on the hardware. The foot restraint positions thus predicted were tested in the Neutral Buoyancy Simulator by astronauts and engineers who determined whether the work they were required to perform was is feasible. These evaluations provided the data for determining the effectiveness of the modelling.

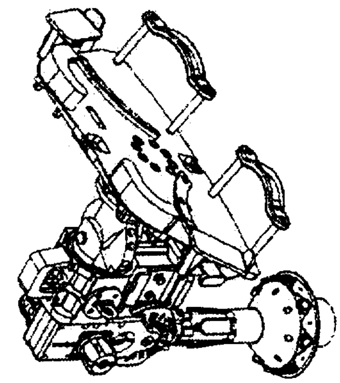


Figure 1. International Space Station type foot restraint. The astronaut places boot toes into the straps on the platform and hooks heels under the ledges on the left side of the drawing (the back of the platform). The joints in the support structure allow many different configurations.

The International Space Station Assembly sequence consists of a series of more than forty hardware launches on Russian and American vehicles, the latter being the U.S. National Space Transportation System, or Space Shuttle. Assembly from components will be achieved through a combination of robotic manipulation and EVA. Robotics will be conducted initially by the Shuttle's robotic arm, the Remote Manipulator System. After the Station becomes habitable, another Canadian-built arm, the Space Station Remote Manipulator System (SSRMS), will be delivered and will take over most Station robotic operations. Assembly Flight 6A, the sixth American launch, will carry the SSRMS and other cargo to the U.S. Laboratory module in January, 1998, The cargo will be attached to a U-shaped pallet which fits in the Shuttle Payload Bay. Marshall Space Flight Center is responsible for packaging the cargo in the pallet, and the design must allow the removal of the hardware by EVA. The astronauts will remove bolts, attach cable connectors, and operate mechanisms in the course of removing the cargo and deploying it to the Station. The majority of these tasks will require the use of appropriately located foot restraints.

METHODS

The human modelling package, Jack, was developed by the University of Pennsylvania as an anthropometric human factors tool. The human model, also referred to as Jack, can be scaled to the standard anthropometric dimensions, allowing the user to develop models that simulate a wide variety of body forms. For these simulations, models for 5'", 50 ᵗʰ, and 95 ᵗʰ percentile humans, by stature, were used, in accordance with NASA standards documentation (NASA, 1995) and published suit data (Pantermeuhl, 1995). Jack attaches kinematics features to any articulated figure. *Rcalistic joint motion limitations were* imposed on the models of the spacesuit, the foot restraint, and the robot. CAD models of the cargo were obtained from the designers: the Canadian Space Agency provided a model of the arm, models of additional pallet cargo came from McDonnell Douglas Aerospace and The Boeing Company, the spacesuit and foot restraint were from Johnson Space Center, and Marshall Space Flight Center supplied models of the pallet and the bracketry used to attach the cargo to it. Versions 5.8 and 5,9 *of Jack* were used in the simulations.

The spacesuited human figure was placed in the appropriate location with respect to the hardware to simulate the task. Reach and visual envelopes were examined, and measurements were taken where appropriate. Once task feasibility was ascertained, the foot restraint was attached to the model's feet and attach points on the pallet were determined. Further data were collected on the foot restraint joint angles, which were read from the software *(Dischinger et al.*, 1996).

After foot restraint attach points were determined, modelling simulations were conducted to attempt to accurately depict body positioning and reach that a spacesuited subject would assume in order to accomplish the tasks required for hardware removal.

An underwater evaluation was performed on mockups built to the design used and generated in the simulations. This test series was conducted at the Neutral Buoyancy Simulator. It consisted of evaluations of the design by eight engineers and astronauts, working in pairs. The subjects attempted the tasks working from foot restraints placed where the modelling simulations predicted they should be and using the joint configurations predicted by the simulations. The tasks were rated according to difficulty. The rating scale indicated whether the design was acceptable (the task could be done by a trained astronaut) or unacceptable, with *varying implications of the latter (minor* modification through complete redesign required). Still photographs were used for comparison with the results of the simulations for correspondence of outcome.

RESULTS

The neutral buoyancy simulations indicated 96% of the tasks should be feasible for trained astronauts of the full range of anthopometric sizes; 94% of the foot restraint locations were rated as acceptable. The percentages *of* predicted foot restraint joint settings which were given acceptable ratings are shown in Table 1.

Configuration variable	Percent acceptable
Orientation at attach point	79
Pitch	59
Roll	78
Yaw	78

Table 1. Percentage of foot restraint *configuration settings predicted by* modelling that were found to be acceptable in test.

Body position models, when compared to photographs of astronauts performing the tasks, were found to be faithful to the real task execution, An example comparison is given in figure 2. In the model and in the actual photograph from the neutral buoyancy simulation, the two astronauts are depicted removing one of the pieces of hardware to be installed *on the Station called the Laboratory* Cradle Assembly (this structure has a grasping latch and will be used for temporary attachment of other hardware to the Station). A model of a fifth percentile woman was generated and used as the figure on the left in 2a. The actual astronaut who performed this task nearly fit this stature category, and the photograph (2b) indicates the modelling closely simulated her reach capabilities.

In *one* case, the attempt to place a foot restraint was unsuccessful; there was no available structure *to which it could be secured.* The task simulation suggested an operational solution to the problem. The task requires that one astronaut pass a long bolt which has been removed from the SSRMS to the other astronaut, who would stow it in a safe place. The only available structure to which the foot

restraint for the second astronaut might be attached put even the largest astronaut out of reach of the hand-off. Based on the modeling results, it was recommended that the recipient be in free float; *i.e.*, tethered to prevent separation from the Station, but not attached to a restraint. When the task was attempted this way in the Neutral Buoyancy Simulator, free-float was found to be an appropriate solution.

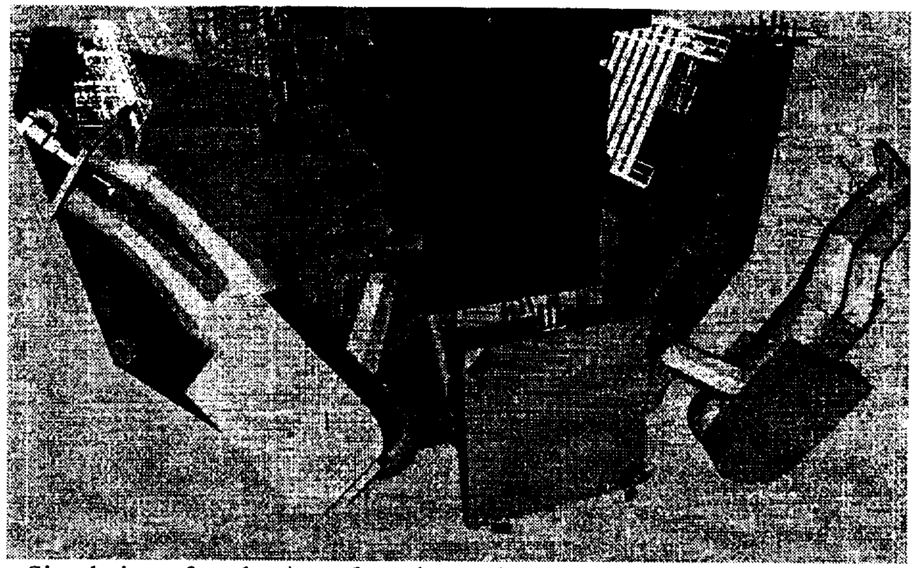


Figure 2a. Simulation of evaluation of reach envelopes required to release cargo from the pallet. The spacesuited model on the left is in a foot restraint attached to the upper left portion of the pallet; she is shown reaching hardware interfaces from this attach point.

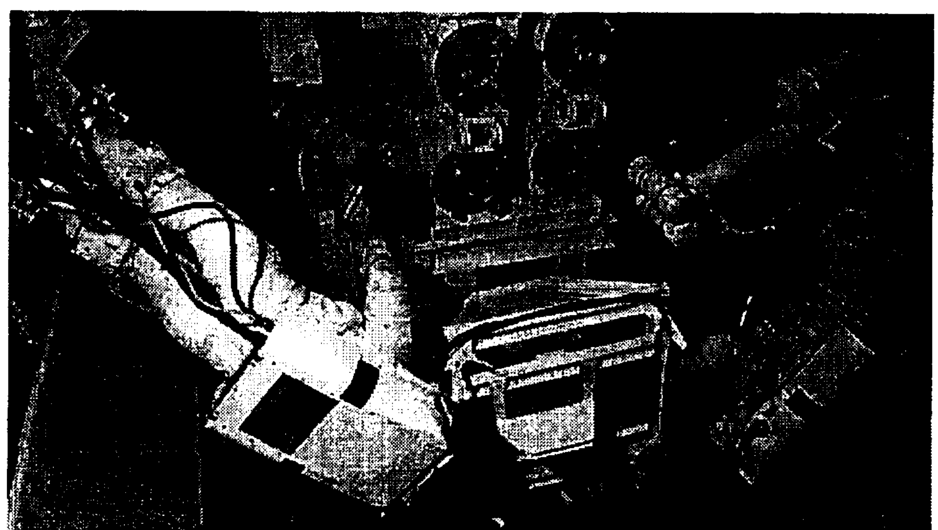


Figure 2b. Performance of the task simulated in 2a in the Neutral Buoyancy Simulator; the task was rated acceptable by astronaut and engineer evaluators, The foot restraint was properly placed, and the body envelopes of the spacesuited figure were appropriate.

DISCUSSION

Jack was found to accurately predict task feasibility for 94 percent of the tasks. It must be noted that this reliability occurred despite considerably lower accuracy in predicting the foot restraint joint settings. These were found (Table 1) to be as low as 59 percent reliable. That is, the suited subjects in the Neutral Buoyancy Simulator found that tasks could be accomplished from the foot restraint placement that had been predicted, but they preferred different orientations. This conflicts slightly with the modelling exercise depicted in Figure 2, which shows concordance

between human figure models and their real test subject counterparts. This demonstrates a limitation of the modelling system, but it is not yet understood which limitations are inherent to Jack and which result from constraints derived from standards documentation. It is clear that at this point, Jack (like other known human models) is incapable of simulating many fine motor tasks. For example, the connectors used in EVA mating of power and data lines are cylinders with bales (levers which engage the two connectors to be joined, pulling them together) on the sides. The EVA astronaut grasps the cylinder along its length and actuates the bale with the thumb, the way one would turn on or off a flashlight. While Jack has grasping capabilities, it cannot yet actuate the bale. Bale actuation is vital to cable connection, and it is thus insufficient to determine that the design provides access to connectors. This is because there are some positions at which one might be able to reach a connector but be unable to slide the bale due to strength or motion limitations. In addition, Jack does not "know" which body positions are comfortable for astronauts, and therefore no such information is available from the modelling. Astronauts, on the other hand, will reposition the foot restraints to compensate for motion limitations not part of the model or for comfort. For the current exercise, these limits on the model capability appear to matter little. Eventually, however, they will likely be accounted for. We are currently planning a neutral buoyancy test which we hope will give us some information about how important information about fine motor constraints is to modelling.

Jack was developed for applications at the earth's surface, Microgravity effects we have included so far are crude. Improvements we plan to attempt include a more realistic neutral body posture and suit-driven motion constraints. The neutral body posture is the somewhat crouched, "arms floating" posture assumed by people in space. The current motion constraints are derived from two- and three-dimensional records of motion studies of suited subjects, That is, they amount to pictures of what subjects are capable of in a suit. A more useful approach is to impose limitations on the suit that are derived from design and pressurization and then to allow the human figure inhabiting it to be so constrained. Heretofore, computational power has limited this capability. The time when this is feasible using workstations is foreseeable.

It should be noted this work does not "validate" Jack as an EVA modelling tool. There are not enough data to conduct statistical analyses on or to draw generalizations from. However, it should give some confidence that this software and other packages like it are of value in understanding gross EVA task design. More to the point, the replacement of neutral buoyancy testing of designs by modelling is not predicted in the near future. While incorporation of fine motor skill constraints is given as a goal, its realization is likely to be in the fairly distant future. A cursory examination of the many fine-tolerance tasks performed on a Hubble Space Telescope upgrade would give even the most enthusiastic modelling supporter pause. This is likewise true of many of the tasks required for Station assembly.

REFERENCES

Dischinger, H. C., Jr., A. M. Evans, and T.E. Loughead 1996. Use of human factors computer modeling to predict locations and settings for crew aids used in extravehicular activity on an International Space Station assembly mission. In Proceedings of the Human Factors and Ergonomics Society, 40th Annual Meeting (Philadelphia, 2-6 Sept.). HFES, Santa Monica, CA, 1254.

Gould, J.D. and C. Lewis. 1985. Designing for usability: key principles and what designers think. Conmmun. ACM. 28('3):300-311 .

National Aeronautics and Space Administration. 1995. ***Man-System integration Standards.*** NASA-STD-3000, Vol. 1, Rev. B. NASA Johnson Space Center, Houston, TX. (July).

Pantermeuhl, J.D. 1995. EMU reach and proximity modeling data. LMES-3 1732, Lockheed Engineering and Sciences Company, Houston.

Page intentionally left blank

URC97036

PACES Participation in Educational Outreach Programs at the University of Texas at El Paso

Rebecca L. Dodge
Pan American Center for Earth & Environmental Studies
University of Texas at El Paso
El Paso, TX 79968
e-mail: dodge @geo.utep.edu

Introduction

The University of Texas at El Paso (UTEP) is involved in several initiatives to improve science education within the El Paso area public schools. These include outreach efforts into the K-12 classrooms; training programs for in-service teachers; and the introduction of a strong science core curricula within the College of Education. The Pan American Center for Earth and Environmental Studies (PACES), a NASA-funded University Research Center, will leverage off the goals of these existing initiatives to provide curriculum support materials at all levels. We will use currently available Mission to Planet Earth (MTPE) materials as well as new materials developed specifically for this region, in an effort to introduce the Earth System Science perspective into these programs. In addition, we are developing curriculum support materials and classes within the Geology and Computer Departments, to provide education in the area of remote sensing and GIS applications at the undergraduate and graduate levels.

Outreach into the Pre-service and In-service Teacher Community

One of our main education goals for 1997 is the development of a multi-year series of workshops for pre-service and in-service teachers in the El Paso area. Public school enrollment in El Paso area schools ranges from 40% to 100% Hispanic. This historically underserved population provides 85% of the students enrolled in the College of Education at UTEP. Of those students enrolled in the College of Education, 85% of the graduates return to area schools to teach. This provides an opportunity for a broad impact by programs introduced to pre-service Education majors at the university, and to ensure the introduction of the MTPE materials into the local public school curricula. It will also generate a population of future education majors familiar with the earth system science perspective. The impact of the continuation of the workshops beyond one year will be multiplied in this region because the students will be exposed to MTPE materials in K-12 and in their education programs at UTEP. In-service teacher participation in these workshops will be limited to no more than 25%, but we expect this participation to add an applied perspective to the workshops because the in-service participants will be carefully recruited from among the top science educators in the El Paso area public school systems.

The El Paso area comprises 12 independent school districts in which technology is an active agent for curricular reform. Technology is being introduced at every grade level, in a variety of applications. The El Paso ISD, one of the city's largest (85 schools), integrates technology in the early years by using computer-based programs to enhance reading/writing skills. Laserdisc technology is evident in every class in the elementary programs. We propose to expand on laserdisc technology use and enhance science curriculum in the classrooms through the introduction of the concepts of earth system science, by using both MTPE materials and materials developed specifically for this region through PACES. Our workshops will focus on providing pre-service teachers with the knowlege of how to incorporate earth system science into their science curricula, and practice in using existing technology, such as laserdiscs, software, and the Internet. in-service participants not only will gain insight into the incorporation of earth system science into their science curricula, but also will mentor the pre-service teachers in cooperative learning groups.

The objectives of' these workshops include:

* Integrate the earth system science perspective and the use of technology into current College of Education science core curricula subjects through the use of MTPE materials.

* Make the whole concept of earth system science relevant by incorporating regional images, both ground and satellite, into the workshops, cooperating with the local NASA University Research Center (PACES) to select and process the satellite images.

* Familiarize workshop participants with available MTPE materials in a hands-on environment.

* Work in cooperative learning groups to develop lesson plans to accompany the MTPE laserdisc and regionally relevent images.

* Put pre-service teachers in contact with designated mentor in-service teachers to use for future referral.

Workshop design:

The Partnership for Teaching Excellence (PETE), is a cooperative project among the Science, Math, and Education departments at UTEP that has fostered a continuity between the science core curricula classes and the teaching methods classes required of BIS (Bachelor of Independent Studies) students. This continuity is created by requiring students to apply the concepts learned in their science/math courses within their educational courses via the teaching methods classes. Our workshops will continue this process by becoming an extension of the teaching methods class, where the students enrolled in the class will be required to attend the workshop as part of their course requirements. The in-service teachers will be chosen from the teacher mentors who are already participating with UTEP in the Urban Systemic Inititave, (USI) which identifies exemplary in-service educators and offers enrichment programs and training. The workshop participants will be placed into several groups, each containing both pre-service teachers and one in-service teacher who will act as a mentor.

Faculty and staff from PACES and the UTEP departments of geology, math, physics, and biology will participate in presentation of materials in their fields, and will serve as resource people for both the BIS students and in-service teachers involved in the workshops. Most of these faculty members are already involved in the PETE program at UTEP and have worked on the development of the core curricula to integrate education methods classes with science and math classes.
The main focus of five of the workshops will be to expose teachers to the concepts and principles of earth system science, utilizing NASA MTPE materials and PACES-developed images. This will include the use of laserdiscs, videos, software, and the Internet. Since laserdisc players are readily available at schools and the laserdisc is a tool that most teachers are familiar with, each of the workshops will include the use of the NASA/CORE Earth Observation laserdisc. To tie in the concepts of earth system science and make those concepts relevent to this area, regional satellite and ground images will also be provided. The format of these workshops will include:

* Introduction to the workshop topic. Introductory material will consist of general information related to the topic and specific information relevant to the local El Paso region. A specially-developed teacher's guide that will contain the introductory material and a listing of MTPE resources will accompany each workshop.

* Hands-on activities within cooperative groups will be used as reinforcement of concepts introduced . These activities will include writing assignments in the form of lesson plans, the use of computers for software demonstration, and using the Internet to obtain updated materials.

* Presentation of NASA materials pertinent to the workshop topic to be used as an introduction, as a reference source for images, and as a demonstration of the use of technology in the classroom. The participants will evaluate MTPE materials available for classroom use at the NASA Regional Teacher Resource Center.

* Presentation, discussion, and evaluation of lesson plans developed by each group during the workshop. During each workshop the participants will employ concepts presented to develop lesson plans related to the NASA and regional materials provided by the workshop conveners. Each group will develop a lesson plan based on one type of technology (i.e., laserdisc, software, video, or Internet) available each workshop and will use a different technology at every workshop to gain experience in all types of technology.

* Preparation of a weekly e-mad journal by all participants that will be reviewed by the co-investigators to evaluate the progress of the workshops and to answer participants' questions. Participants will also keep an electronic portfolio with a compilation of their work. These portfolios will also be used by workshop conveners as an evaluation tool.

Workshop Topics: Earth System Science

Five workshops will present the earth system science perspective using MTPE and UTEP materials. Each workshop will focus on one aspect of earth system science: an introductory workshop, the lithosphere, the biosphere, the atmosphere and hydrosphere, and astronomy. A sixth, and final, workshop will consist of a field trip to the NASA Regional Teacher Resource Center maintained by the New Mexico Space Grant Consortium located at New Mexico State University in Las Cruces, New Mexico. The El Paso area satellite images processed by PACES and the lesson plans developed by the workshop participants will be available via the PACES web site.

College-level Education

During 1997 PACES will offer an upper-level undergraduate class in Remote Sensing Applications within the Geology Department. This is designed to be primarily a hands-on laboratory-type class, using the existing hardware, software, and imagery resources of PACES and the Geology Department, to teach students the about the practical image processing, image interpretation, and mapping/GIS applications that will enable them to use satellite imagery in their future graduate work and on the job. We will rely heavily on imagery from within the PACES area of geographic interest, so that field-checking will be an important aspect of the class. The course outline for this class has been condensed into a workshop for this conference. A graduate-level class in image processing and GIS is in the planning stages.

PACES personnel are working closely with professors from the Geology and Computer Departments to identify imagery and applications that can be used within existing classes. For example, PACES will be processing imagery that can be used to demonstrate basic principles in both Physical and Historical Geology classes. We are also working with professors to develop image-based exercises for Structural Geology, Geomorphology, Sedimentology , and Field Camp classes.

PACES personnel are also actively working with professors in several on-campus departments, to identify the appropriate imagery and technology that can be used to introduce MTPE materials within disciplines outside of the Geology and the Computer Departments. For example, PACES is processing a set of "change-detection" imagery of the El Paso area, spanning almost 20 years, which can be used for a variety of applications. We are coordinating these activities with the

Center for Environmental Resource Management (CERM) here at UTEP, in order to reach the broadest audience within the University.

Acknowledgement

We wish to thank NASA for its support of this project through the grant NCCW-0089.

URC97037

52

Current Methods of Uncertainty Analysis in Human Health Risk Assessment

Sunil Donald[1] and Timothy J. Ross[2]

Abstract

This paper discusses the methods currently used for uncertainty analysis in human health risk assessment. In a broader sense this uncertainty analysis can be understood in the context of any general engineering system. Since uncertainty arises because of the lack of exact knowledge about truth, it pervades our understanding of all systems, and has been the cause of a number of unexpected outcomes when making predictions about such systems. In the field of artificial intelligence, uncertainty is segmented into three dimensions: 1) uncertainty due to randomness; 2) uncertainty due to vagueness; and 3) uncertainty due to ambiguity. The three powerful methods of probability theory, evidence theory and fuzzy set theory are used in the discussion of this important issue.

1. Introduction

Risk assessment, a process which enables one to organize and systematically analyze diverse scientific information in a single framework, has gained much popularity in the recent years. Since the passage of the Clean Air Act in 1970, the number of environmental regulations has increased tremendously. The economic burden of the increasing number of regulations has prompted federal agencies to weigh the costs of regulations against benefits realized from them. As an important tool in determining the benefits part of the cost-benefit equation risk assessment is being scrutinized closely for it's validity and effectiveness. In a recent directive, Congress charged NRC to evaluate the risk assessment methods used by the EPA to set the standards for 189 air pollutants. Amongst the issues identified by NRC as requiring revisions, the EPA was asked to lay more importance on the uncertainty and variability issues. The document questions the EPA claim that quantitative uncertainty assessment is usually not practical or necessary for site risk assessments. The committee discourages the use of single point estimates, and strongly recommends that the agency use representations which are apt to include all the sources of error and also encourages that, "*to the greatest extent feasible, EPA should present quantitative as opposed to a qualitative representations of uncertainty*" [NRC, 1994]. Due to Imitations on financial, technological and temporal resources, the need to consider all the relevant information and therefore conduct a formal uncertainty analysis has become very crucial. This has inspired a significant growth in the research on uncertainty analysis. In attempts to reduce uncertainty, the context of models used in risk assessment is shifting from simplistic statistical models yielding point estimates to more comprehensive distributional models. Currently most of the uncertainty models being developed are based on probability theory. Though belief and plausibility measures, possibility measures and fuzzy set theory have made significant contributions in the artificial intelligence (AI) community, they have had restricted or no exposure to the risk assessment community. With the acknowledgment that expert judgments play an important role in risk assessment, techniques used mainly in AI are trickling into risk

[1] PhD Candidate, Dept. of Civil Engineering, Univ. of New Mexico, Albuquerque, NM, 87131
[2] Regents' Professor, Dept. of Civil Engineering, Univ. of New Mexico, Albuqeruque, NM 87131

assessments. This paper attempts to reflect upon the methods currently used for uncertainty analysis in human health risk assessment.

2. Definition of Uncertainty

Uncertainty, the result of the lack of exact knowledge about the truth, pervades all analysis and systems, and has been the cause of a number of unexpected outcomes. The definition of uncertainty and its roots has been construed by many authors in various ways. In the field of artificial intelligence, uncertainty is segmented into three dimensions: 1) uncertainty due to randomness; 2) uncertainty due to vagueness; and 3) uncertainty due to ambiguity. Randomness exists when evidence points to the occurrence of conflicting events, that is any of the events can occur with a given probability. On the other hand, ambiguity arises due to the indecision in assignment of an element to a given set and vagueness is due to the non-distinct boundaries of the set. Uncertainty only due to vagueness is usually best modeled by fuzzy set theory, while evidence theory and possibility theory can be used to represent random, ambiguous and vague uncertainty. It is broadly accepted in the risk assessment community that, on the basic level, uncertainty can be classified into three main groups: 1) randomness; 2) bias; and 3) variability. In a few papers uncertainty has been specifically distinguished from variability in that, uncertainty is identified as arising from the imperfection in knowledge about the true value, while variability is the real variation among individuals [Hattis and Burmaster, 1994]. Thus uncertainty can be reduced by the collection of more data, while real variability will remain constant, although it may be more accurately known with the acquisition of more data.

Bias, which is a result of the systematic errors in measurements, occurs more in subjective judgments due to the reliance on judgmental of representativeness, availability, adjustment and anchoring heuristics. Representativeness is a heuristic used when a subject is asked to judge the probability that a specific event belongs to a certain class or process; availability is often employed when the expert is asked to determine the frequency of a class or the plausibility of a certain development; and, adjustment is used in numerical predictions when a relevant value is available to the subject.

3.0 Methods of Uncertainty Analysis

Three main theories have been put forth to represent the uncertainty in the data. These theories are distinguished according to the underlying concept of representing various forms of uncertainty. Probabilistic theories are used to represent the outcome of random events and are based on the concept of chance. Within the probability there are two groups, the objectivists who adhere to the frequentist concepts and the subjectivists who subscribe to the Bayesian concepts. Both these theories are subjected to the additivity axiom, on the other hand, non-probabilistic theories such as fuzzy measure theory which includes evidence and possibility theories, do not require a strict adherence to this axiom. Finally, uncertainty arising from vagueness of set boundaries, as in a set of numbers representing linguistic variables (such as low, medium, high) is most appropriately modeled using fuzzy set theory.

3.1 Probability Theories - Frequentist and Bayesian

In the past, uncertainty was usually associated with randomness of the system behavior and hence probability has been the basis for defining uncertainty. A probability space is defined on a domain Θ (also called the universe of discourse) and the algebra of sets A, and is given by the probability triple (Θ, A, P), where P represents the probability of the occurrence of A on Θ. Probability in the classic sense (Frequentist and Bayesian) is assumed to satisfy the following Kolmogorov's axioms:

Axiom 1: $0 \leq P(A) \leq 1$

Axiom 2: $P(\varnothing) = 0$; $P(\Theta) = 1$

Axiom 3: $P(A) + P(\overline{A}) = 1$

The third axiom is what distinguishes probability theory from non-probabilistic theories. Within probability theory there are two schools of thought, the frequentist and the Bayesian. Frequentists consider probability as the relative number of occurrences of a given event over the total number of trials. Given a large enough sample, N, and if N(A) is the number of occurrences of event A, then, $P(A) = \frac{N(A)}{N}$, represents the probability of event A. As opposed to frequentists, Bayesians believe that the probabilities are degrees of belief deliberately constructed and adopted on the basis of all the available evidence, both objective and subjective. Bayesian method is based on the celebrated Bayes theorem of conditionalization:

$$P(H|e) = \frac{P(e|H)P(H)}{P(e)}$$

stating that the belief accorded to a particular hypothesis H upon obtaining evidence e can be computed by multiplying one's prior belief $P(H)$ and the likelihood that e will materialize assuming H is true.

Due to the familiarity and objectivity associated with frequentist analysis, most of the analysis in risk assessments use frequentist methods. These methods are extensively used in the derivation of dose-response relationships for carcinogens, reference doses for non-carcinogens, and the determination of concentration profiles at various hazardous waste sites [EPA, 1989]. Models such as, probit, multi-hit, linear and linear multistage, all reflect the use of frequentist methods. Bayesian techniques have been used to a limited extent by EPAs office of air to determine non-cancer health effects of air pollutants [Whitfield and Wallsten, 1989].

Debates over which uncertainty analysis method are most appropriate has usually centered around the usage of probability distributions as opposed to point estimates (such as mean or 95th percentile values). The trend however seems to be in the direction of using the entire probability distributions and using Monte Carlo simulation to propagate the distributions. Smith [1994] conducted a health risk assessment of individuals exposed to volatile solvents by drinking water ingestion and inhalation using this approach. Hoffman and Hammonds [1994] demonstrate the use of Monte Carlo approach to propagate combined uncertainty and variability (referred to as type A and

type B uncertainty). The main drawbacks of a Monte Carlo approach are its inability to support sensitivity analysis, high computational effort and the dependence on the accuracy of the probability y distributions. These drawbacks, however, are being overcome by efficient sampling techniques and the inclusion of sensitivity analysis techniques into the softwares. Methods of selecting the appropriate distributions for all the variables has been studied in depth by Seiler and Alvarez [1996].

In addition to conventional Monte Carlo simulations, fuzzy arithmetic has been added to the sampling scheme for the simulation of fishery population [Ferson, 1993]. In this method, samples are taken as fuzzy numbers and then combined using interval analysis techniques.

3.2 Non-Probabilistic Theories

Amongst the non-probabilistic theories, evidence theory and fuzzy set theory are the only theories that are currently being used in risk analysis [Krause and Fox, 1995; Lein, 1992; Bardossy *et al.*, 1993].

3.2.1 *Evidence theory*

Evidence theory (also known as Dempster-Shafer Theory), originally conceived by Arthur Dempster [Dempster, 1968] to determine the lower and upper probabilities, and later extended by Glenn Shafer to embody epistemic degrees of belief, comprises of assigning probabilities on the basis of the belief in evidence obtained to support a proposition. Therefore, belief is assigned to individual pieces of evidence, and not considered to yield total probabilities directly. In this theory the elements of a finite set and its subsets are identified and then beliefs are assigned to each one of these sets. The superset Θ of all the subsets A is usually referred to as the frame of discernment and the subsets for which the assigned belief is non-zero are called the focal elements. Beliefs are assigned not only to the subsets, but also to all the individual elements of the hypothesis sets. If Θ is the frame of discernment, then a function $m 2^{\Theta}$ + [0,1] is called a basic probability assignment when,

1. $m(\varnothing) = 0$, and
2. $\sum_{A \subset \Theta} m(A) = 1$

The basic probability assignment *m(A)* represents the belief one commits exactly to *A*, and not the total belief committed to *A*, i.e., not to any other subsets of Θ. The total belief committed to *A* is determined by the following equation,

$$Bel(A) = \sum_{A_i \subset A} m(A_i)$$

Now, if *Bel* $(\overline{A})$ represents the extent to which one believes in the negation of *A*, then the dual of a *belief measure* called *the plausibility measure, Pl (A), is* given as,

$$Pi(A) = 1 - Bel(\overline{A}) = \sum_{A_i \cap A \neq 0} m(A_i)$$

If@ is a finite set, and 2^{Θ} represents the set of all subsets of Θ, then a belief function is supposed to satisfy the following axioms [Shafer, 1976]:

1. $Bel(\emptyset) = 0$

2. $Bel(\Theta) = 1$

3. For every positive integer n and every collection $A_1, \ldots, A_n$ of subsets of Θ,
$Bel(A_1 \cup \ldots \cup A_n) \geq \sum_i Bel(A_i) - \sum_{i<j} Bel(A_i \cap A_j) \ldots\ldots + (-1)^{n+1} Bel(A_1 \cap \ldots \cap A_n)$

From the *axiom* 3 of the belief functions it can be derived that,

$Bel(A) + Bel(\overline{A}) \leq 1$, and

$Pl(A) + Pl(\overline{A}) \geq 1$

The above equations represent the fundamental difference between the evidence theory and Bayesian theories. The difference between 1 and *(Bel(A)+ Bel* ($\overline{A}$) *)* represents the ignorance due to the lack of knowledge. Though evidence theory offers an excellent framework for modeling expert judgments, it's use in human health risk assessments has been very limited. Krause et al. [1995] explore the idea of using belief and plausibility measures in deriving plausible conclusions about the risk from carcinogenic chemicals and for combining qualitative and quantitative measures in risk characterizations.

3.2.2 *Fuzzy Set Theory*

Fuzzy sets, introduced by Zadeh in 1965 [Zadeh, 1965], have been extensively used in control, optimization and decision making systems. Fuzzy logic, in comparison to the classical logic exploits the tolerance in systems and allows for partial truths to be assigned to various states. Instead of the truth values being in either the O or the 1 state, a partial truth within the interval [0, 1] is used. A fuzzy set, $\tilde{A}$ is given by a membership function, $\mu_{\tilde{A}}$ (x), on a real line [0,1], and represents the degree to which an element, x belongs to the set $\tilde{A}$. Hence, by allowing a range of numbers a fuzzy model most appropriately models vague human thinking. Membership functions and probability distribution functions should not be confused with each other, as one deals with ambiguity in assignment of an element to a particular set, while the latter deals with randomness in the occurrence of an event. An excellent description of the methods for derivation of membership functions and propagation of such functions is given in Ross [1995].

There has been very limited use of fuzzy set theory in risk assessments. Bardossy *et al. [1993]* use fuzzy non-linear regression analysis in the determination of dose-response relationships of N-nitroso compounds. In this model, the equation for dose response relations consists of parameters which are given by a range determined by fuzzy non-linear regression. There have been more applications of fuzzy set theory in the risk management of hazardous waste sites [Donald and Ross; 1996].

4. Conclusions

Reduction and appropriate characterization of uncertainties in risk assessment has been the main objective of many risk assessors. As far as EPA's documents are concerned, uncertainties are mostly represented qualitatively, and not quantitatively. EPA requires that the 95th percentile values for carcinogenic substances and a safety factor in the

orders of magnitude of 10 for non-carcinogenic substances be used in the risk assessment model. These methods rely on point estimates and fail to capture the entire range of uncertainty. Most papers written on modeling uncertainty in risk assessment deal with the propagation of numerical values of uncertainties derived from probabilistic concepts. Monte Carlo simulation has been the most widely used method for the propagation of uncertainties. A recent paper by Krause and Fox [1995], points out the implications of subjective judgments in risk assessments and lays a framework for inclusion of belief and plausibility measures. A few fuzzy logic models, such as regression analysis to obtain dose-response relationships have been developed in the recent past. Much research still needs to be done to combine probabilistic and non-probabilistic concepts for including subjective as well as objective information.

References

1. Bardossy, A., Bogardi, I. and Duckstein, L., "Fuzzy Nonlinear Regression analysis of Dose-Response Relationships," ***European Journal of Operational Research 66,36-51,*** 1993.
2. Dempster, A. P., "Upper and Lower Probabilities Induced by a Multivalued Mapping," Ann. Math. Stat. 38, 325-339, 1967.
3. Donald, S, and Ross, T. J. "Use of Fuzzy Logic and Similarity Measures in the Risk Management of Hazardous Waste Sites," ***3rd Congress on Computation in Civil Engineering,*** Anaheim, CA, 1996.
4. Ferson, S., "Using Fuzzy Arithmetic in Monte Carlo Simulation of Fishery Populations," ***Proc. of the Int. Symposium on Management Strategies of Exploited Fish Populations,*** Alaska Sea Grant College Program, AK-SG-93-02, 595-608, 1993.
5. Hattis, D. and Burmaster, D. E. "Assessment of Variability and Uncertainty Distributions for Practical Risk Analysis," Risk Analysis 14(5), 713-730, 1994.
6. Hoffman, O. F. and Hammonds, J. S., "Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability," ***Risk Analysis 14, 707-712, 1994.***
7. Krause, P., Fox, J., and Judson, P, "Is there a Role for Qualitative Risk Assessment?" ***Proc. of the Eleventh Conf. on Uncertainty in Artificial Intelligence,*** Eds. Besnard Phillippe and Steve Hanks, Morgan Kaufmann Publishers Inc., 1995.
8. Lein, J. K. " Expressing Environmental Risk Using Fuzzy Variables: A Preliminary Examination," The Environmental Professional 14,257-267, 1992.
9. National Research Council, "Science and Judgment in Risk Assessment," National Academy Press, 1994.
10. Ross, T. J., ***Fuzzy Logic With Engineering Applications,*** McGraw Hill, Inc., 1995.
11. Seiler. F. A. and Alvarez, J. L. "On the Selection of Distributions for Stochastic Variables," Risk Analysis 16 (l), 5-18, 1996.
12. Shafer, G., ***A Mathematical Theory of Evidence,*** Princeton University Press, 1976.
13. Smith, R., "Use of Monte Carlo Simulation for Human Exposure Assessment at a Superfund Site," ***Risk Analysis*** 14, ***433-439,*** 1994.
14. Thompson, K. M., Burmaster, D. E. and Crouch, A. C. "Monte Carlo Techniques for Quantitative Uncertainty Analysis in Public Health Risk Assessments," Risk Analysis, 12 (1), 53-63, 1992.
15. U. S. Environmental Protection Agency (EPA), "Risk Assessment Guidance for Superfund," Vol. 1, Health Evaluation Manual (Part A), Interim Final, EPA 540/1 -89-002, 1989.
16. Whitfied, R. G. and Wallsten, T. S. "A Risk Assessment for Selected Lead-Induced Health Effects: An Example of a General Methodology," ***Risk Analysis 9, 197-208, 1989.***
17. Zadeh, L. A., "Fuzzy Sets," ***Information control 8, 338-353,*** 1965.

Acknowledgement

This research was supported in part by an appointment to Student Research Participation Program at the U. S. Army Environmental Policy Institute administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U. S. Department of Energy and USAEPI.

URC97038

~~3~~/61

Hybrid Co-evolutionary Motion Planning via Visibility-Based Repair

Gerry Dozier Shaun McCullough Edward Brown Jr. Abdollah Homaifar Mar-wan Bikdash

NASA Center for Autonomous Control Engineering
North Carolina A&T State University
Greensboro, NC 27411, USA
gvdozier@ncat.edu

Abstract

This paper introduces a hybrid co-evolutionary system for global motion planning within unstructured environments. This system combines the concept of co-evolutionary search along with a concept that we refer to as the visibility-based repair to form a hybrid which quickly transforms infeasible motions into feasible ones. Also, this system makes use of a novel representation scheme for the obstacles within an environment. Our hybrid evolutionary system differs from other evolutionary motion planners in that (1) more emphasis is placed on repairing infeasible motions to develop feasible motions rather than using simulated evolution exclusively as a means of discovering feasible motions, (2) a continuous map of the environment is used rather than a discretized map, and (3) it develops global motion plans for multiple mobile destinations by co-evolving populations of sub-global motion plans. In this paper, we demonstrate the effectiveness of this system by using it to solve two challenging motion planning problems where multiple targets try to move away from a point robot,

KEYWORDS: Evolutionary Algorithms, Co-evolution, Motion Planning, Visibility Graph, Visiblity-Based Repair, Radcliffe's Crossover, Seed Crossover, Seed Mutation, Global Motion Planning, Sub-Global Motion Planning

1 Introduction

Evolutionary Algorithms (EAs) are search methods that evolve a population of candidate solutions through the use of natural selection. EAs typically solve difficult problems for which traditional search paradigms yield unsatisfactory results. EAs have been successfully applied to a variety of areas such as design optimization, machine learning, constraint satisfaction, and constrained optimization [9]. Recently. there has been a growing number of successful applications of EAs to the area of motion planning [2 3, 6, 9, 11]. The Motion Planning Problem [5, 7] can be stated as follows. Given an environment $E(R, X, T, O)$ where R represents some robot, X represents the starting point (or point of origin), T represents the goal or dest inat ion point and O represents a set of obst acles, find a collision free (feasible) path from X to T (path planning phase) that R can traverse (navigation phase), Many of the evolutionary motion planning systems rely on simulated evolution almost exclusively as a means of discovering a feasible motion for R.

This paper introduces a hybrid co-evolutionary system for global motion planning within unstructured environments. This system combines the concept of co-evolutionary search along with a concept that we refer to as the visibility-based repair to form a hybrid which quickly transforms infeasible motions into feasible ones. Also, this system makes use of a novel representation scheme for the obstacles within an environment. Our hybrid evolutional y system differs from other evolutionary motion planners in that (1) more emphasis is placed on repairing infeasible motions to develop feasible ones rather than using simulated evolution exclusively as a means of discovering feasible paths, (2) a continuous map[1] of the environment is used rather than a discretized map, and (3) it develops global motion plans for multiple mobile destinations by co-evolving populations of sub-global motion plans. In this paper, we demonstrate

[1] [6, 9] also use use continuous maps

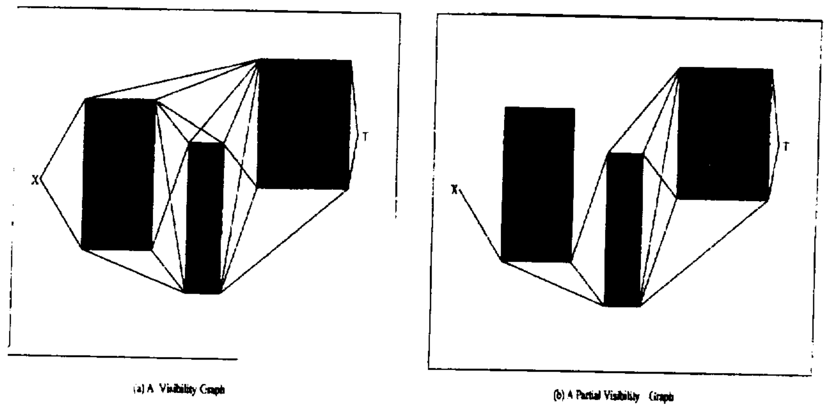


Figure 1: Full and Partial Visibility Graphs

the effectiveness of this new hybrid system by using it to solve two challenging motion planning problems where multiple targets try to move away from a point robot.

The remainder of the paper is organized as follows. Section 2 provides a brief introduction to the concept of visibility-based search. In Section 3, we present our hybrid co-evolutionary system called GEPOA-II (a Global Evolutionary Planning and Obstacle Avoidance system) in detail and, in Section 4, reintroduce our test suite. In Section 5, we present our results and conclusions and, infection 6, we discuss some directions for future research.

2 Visibility-Based Search

A visibility-based search algorithm can be regarded as any search procedure that uses a visibility graph (VG)[3, 5] to aid in the discovery of feasible paths. A VG is a graph, (V, E), where V *is* the set of all vertices of the obstacles within an environment including the coordinates of the robot (or starting position) and the destination, and E is the set of all edges connecting any two vertices in V that do not pass through any obstacles within the environment.

Figure la shows an example of a visibility graph. Notice that X is connected only to the vertices that are reachable by way of a straight-line segment that does not cut or pass through (or violate) any obstacles. Notice also that the vertices which are visible to X are connected, in similar fashion! to other vertices that are visible to them (again by way of straight-line segments). Once a VG has been constructed for a given environment, usually an A^* search algorithm is used to find the shortest path between starting and destination points.

It is not always necessary to construct a complete VG for an environment. Some researchers have experienced a great deal of success with using partial VGs(PVGs) [3]. Figure lb shows an example of a PVG. An advantage of using a PVG rather than a VG is that a PVG requires less computation. One disadvantage of using a PVG is that it may not contain an optimal path. Both of these methods perform poorly on path planning problems with dynamic environments.

3 GEPOA-II

GEPOA-II is a co-evolutionary version of a successful hybrid evolutionary planner named GEPOA-I[4]. We developed GEPOA-II in an effort to take advantage of the decompositional nature of global motion planning. In GEPOA-II, global motion planning problems are decomposed into smaller, sub-global problems. A number of EAs are then used to quickly develop sub-global motion plans (one EA for each sub-problem) which are combined to represent a global motion plan. In this section, we discuss seven salient attributes of GEPOA-II. These attributes are as follows: the representation of environments, the concept, of visibilit y-based repair, the representation of candidate paths (CPs), the visibility-based repair algorithm, the evaluation funct ion, the selection algorithm, and the evolutionary operators. We conclude this section by providing the parameter settings for three GEPOA-II hybrids that will be tested with a test suite of two path planning problems.

3.1 Representation of an Environment

An obstacle within an environment is represented as a set of $\lceil \frac{k_i}{2} \rceil$ intersecting line segments, where k_i represents the number of vert ices of obst acle i. Each line segment, $l_{i,j}$, connects two distinct vertices $v(i, j)$

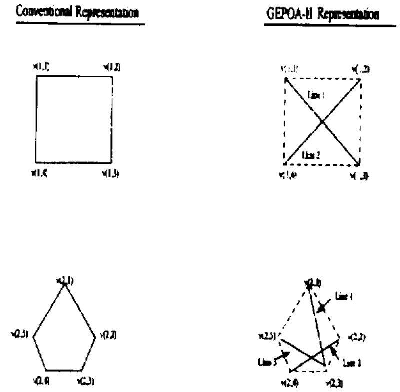


Figure 2: obstacle Representation in GEPOA-II

and $v(i, j + \lfloor \frac{k_i}{2} \rfloor)$ (where $j \leq \lceil \frac{k_i}{2} \rceil$) of obstacle i. Associated with each vertex within the environment, is a value which 'represents the number of obstacles that the vertex is contained by. This value is referred to as the 'containment value' (CV) of a vertex. If a vertex lies along the boundary of an environment its CV is assigned a value of ∞.

Figure 2 provides an example of how obstacles are represented in GEPOA-II. Notice, in Figure 2, that the four sided obstacle (Obstacle 1) is represented by only two lines in GEPOA-II. Line 1 of Obstacle 1 connects vertices $v(1, 1)$ and $v(1, 3)$ while Line 2 connects vertices $v(1, 2)$ and $v(1, 4)$. Obstacle 2 has five sides and is represented in GEPOA-II using three lines. Line 1 connects $v(2, 1)$ and $v(2, 3)$, Line 2 connects $v(2, 2)$ and $v(2, 4)$, and Line 3 connects $v(2, 3)$ and $v(2, 5)$. Since each of the vertices are contained by only one obstacle, the CV for each vertex is 1.

3.2 Visibility-Based Repair

Visibility-based repair (VBR) is performed as follows. When an obstacle, o_i, lies along a straight-line segment between two nodes P and Q, each line of o_i is checked to see if it is intersected by PQ. If a line of o_i is intersected by PQ then a repair node is created using the following set of rules:

Rule 1: if the CVs of a line's vertices are both equal to one, then the repair node is selected to be a point along an extention[2] of the vertex which is closer to the point of intersection;

Rule 2: if the CVs of a line's vertices are different, then the repair node is selected to be a point just outside of the vertex which has the lower CV;

Rule 3: if the CVs of a line's vertices are greater than one and equal, then the repair node is selected to be a point just outside of the vertex which is farther from the point of intersection.

Figure 3 shows an example of how VBR can be used to transform an infeasible path into one that is feasible. In Figure 3a, an infeasible path XPT *is* shown. The path XPT *is* infeasible because the line segment XP passes through Obstacle 1 and the line segment PT passes through Obstacle 3. Before proceeding further, notice that each vertex in the environment shown in Figure 3a has a CV of one.

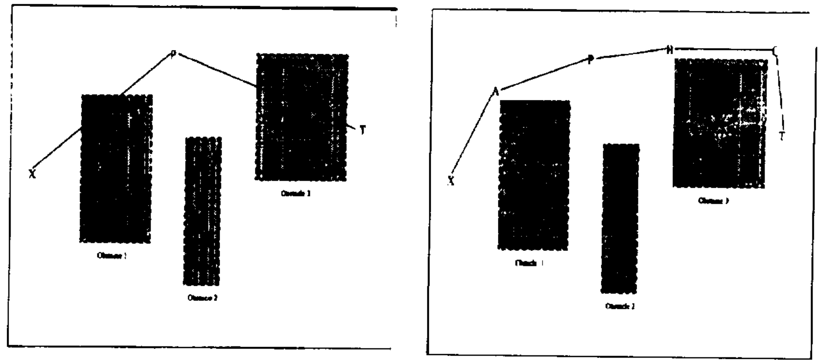
Using VBR, the line segment XP can be repaired to XAP. *Since* XP intersects Line 1 of Obstacle 1, a repair node corresponding to a point just outside of either v(1,1) or v(1,3) must be selected. By applying Rule 1, Node *A*, which corresponds to a point just outside vertex v(1, 1), is selected as the repair node.

Similarly, the line segment PT can be repaired to $PBCT$. Again Rule 1 must be applied to Line 1 and Line 2 of Obstacle 3. The repair node that results from the intersection of PT and Line 1 is Node *B*. The repair node that results from the intersection of PT and Line 2 is Node C". Figure 3b shows the result of using VBR on XPT. The repaired, feasible version of XPT is $XAPBCT$.

3.3 Representation of Candidate Paths

An individual representing a candidate path (CP) contains two fields. The first field is a chromosome which contains a gene corresponding to the cartesian coordinates of each node of the path (where each node of a path is connect by a straight-line segment). The second field is called the seed. The seed of an individual is the gene that will be crossed or mutated to created an offspring. Initially, an individual will have only three genes: the start gene, the seed gene and the destination gene. Repair genes are inserted

[2] The distance outside of an obstacle at which a repair node is placed is a user specified parameter.



a) Before Visibility-Based Repair b) After Visibility-Based Repair

Figure 3: Visibility-Based Repair

into the chromosome by the visibility-based repair algorithm each time a straight-line segment of an individual is found to pass through an obstacle. The third field is a value referred to as the violation distance. The 'violation distance' represents the euclidean distance of the CP which cuts through one or more obstacles. The fourth field records the euclidean dist ante of the path from the start to dest inat ion.

3.4 The VBR Algorithm Used by GEPOA-II

Given a CP, the VBR algorithm used by GEPOA-II works as follows. Each obstacle within the environment is checked with each straight line segment from the start gene to the destination gene of the CP until a segment is found that passes through the obstacle. The infeasible segment is repaired via VBR and the process is repeated using the next obstacle.

For an example of how this repair algorithm works, notice once again Figure 3. When given the path XPT the algorithm works as follows. Obstacle 1 is checked to see if it is violated by segment XP. Since it is, a repair gene (Node A) *is* generated and Obstacle 2 is then considered. Obstacle 2 is checked to see if it is 'cut' by segment XA. Since it is not 'cut' by segment XA, Obstacle 2 is checked with segment AP then segment PT. Since there are no more segments to inspect, Obstacle 3 is considered. Obstacle 3 is checked to see if it is 'cut' by segments XA, and AP. Finally, Obstacle 3 is checked to see if it is 'cut' by PT. Since it is, two repair genes are generated (Nodes B and C) and the algorithm terminates.

3.5 Evaluation and Selection

The evaluation function computes the euclidean distance of each straight line segment of the path that an individual represents as well as the euclidean dist ante of each segment oft he path that passes through one or more obstacles, called the violation dist ante. GEPOA-II uses a modified version of tournament selection, with a tournament size of 2, to select individuals to become parents. The selection process is as follows. Two individuals are randomly selected from the current population. If the violation distances of the two are different then the individual with the smaller violation distance is selected to be *a* parent. If the violation distances are the same then the individual with the smaller 'overall' distance is selected.

3.6 The Evolutionary Operators

GEPOA-II uses three operators to create and/or refine individuals. The first operator, the VBR algorithm, is applied to parents representing infeasible CPs 25% of the time as well as all newly created offspring. The other two operators, (1) a version of Radcliffe's Crossover [10] that we refer to as seed crossover and (2) a version of uniform mutation we refer to as uniform seed mutation, are applied only to feasible CPs.

Seed crossover is as follows. Given two seed genes $s_1 = (x_1, y_1)$ and $s_2 = (x_2, y_2)$, a seed gene for an offspring, $s_{off} = (rnd(x_1, x_2) + N(0, 4.0), rnd(y_1, y_2) + N(0, 4.0))$, is created where rnd is a uniform random number generator and N (O, 4.0) is a gaussian random number with zero mean and a standard deviation of 4.0. The resulting offspring, (X, s_{of}j, T), has a chromosome containing three genes: a gene corresponding to the start node, the seed node, and the goal node. The offspring then undergoes VBR and may have additional repair genes added by the VBR algorithm.

In uniform seed mutation, either the x or y coordinate of a parent is mutated using uniform mutation to create a seed gene for an offspring. A resultant offspring created by seed mut at ion is similar to one created by seed crossover in that it also has a chromosome cent aining three genes. Once again the offspring undergoes VBR and may have additional repair genes added by the VBR algorithm.

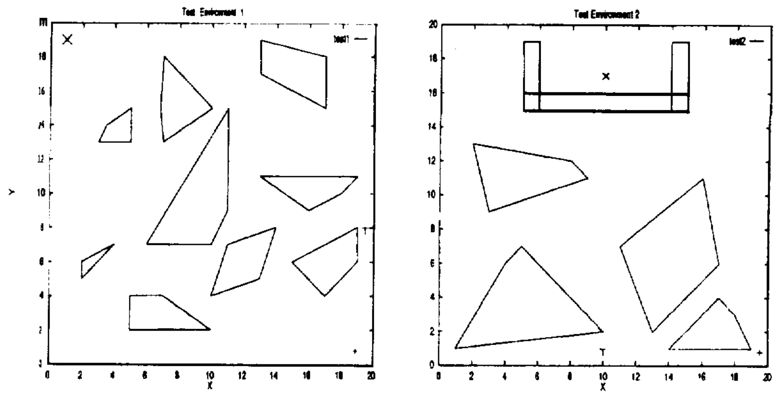


Figure 4: The Two Test Environments

3.7 Attribute Settings for Three Co-evolutionary Hybrids

Three GEPOA-II co-evolutionary hybrids with population sizes of 5, 10, and 20 were tested with a test suite of two motion planning problems. These three hybrids differ only in the size of the two populations that they co-evolve. The first of the two populations co-evolved by these hybrids contains motions from the point robot to the first target. The second population contains motions from the first target to the second target. The hybrids randomly generate their initial populations, use a seed crossover rate of 0.5 and a uniform seed mutation rate of 0.5. For each generation, only one offspring is create. This offspring replaces the worst fit individual in the population.

Every ten generations R *is* advanced a maximum of one unit along the shortest path to T_1 developed by the system. The amount of advancement is based on the change of direction of R. If no change of direction is needed, R *is* advanced one unit. If the change of direction is 90° or greater, R is not advanced. After R has been advanced, T_1 *is* allowed to move a half unit in a direction (north, south, east, or west) furthest away from the current position of R, and T_2 is allowed to move a half unit in a direction furthest away from T_1'. This process is repeated until R reaches T_1. At this point, the second population is used exclusively to navigate R to T_2.

4 The Test Suite

Figure 4 shows our test suite of two motion planning problems that will be solved by the three hybrids described above. In each of the test environments, X represents the starting point, T represents the first mobile destination, and "+" represents the second mobile destination. In Test Environment 1, T is located at (8.0,19.5) and "+"is located at (1.0,19.0). In Test Environment 2, T is located at (1. 0,10.0) while "+" is located at (1.0,19.5).

5 Results and Conclusions

Each hybrid was run 50 times on each of the test environments. Each run was alotted a maximum of 200 moves. A run was considered successful if the robot reached both targets. On each run, the hybrids never failed to find a feasible solution within the initial population. Also, the motion plans evolved by the hybrids allowed the robot to reach the target 100% of the time. This is an indication of how effective visibility-based repair can be.

The results of the performances of the hybrids on both of the test environments are presented in Figures 5 an 6. They are organized into a matrix where the columns (from left to right) correspond to:

- the population size of the hybrid (P),
- the average length of the first feasible solution found $\overline{(Ln.1^{st})}$ from R to T to +,
- the standard deviation of the length of the fist feasible solution found during each successful run ($\sigma(Ln.1^{st})$),
- the average number of moves needed to reach the first target ($\overline{Moves1}$), and
- the standard deviation of the number of "moves need to reach the first target ($\sigma(Moves1)$).
- the average of the total number of moves needed to reach the first and second target ($\overline{\text{Total}}$), and
- the standard deviation of the total number of moves needed to reach the first and second targets ($\sigma(Total)$).

In Figure 5, one can see that as the larger the population size the better the performance. This is especially the case when viewing the $\overline{Total}$ column. However larger population sizes require more

[3] Each time R, T_1, or T_2 is moved each individual in the population is re-evaluated.

P	$Ln.1^{st}$	$\sigma(Ln.1^{st})$	$Moues\ 1$	$\sigma(Moves1)$	$Total$	$\sigma(Total)$
5	34.125	3.250	34.960	7.665	73.260	14.574
10	32.718	1.844	33.980	7.928	68.980	13.568
20	31..746	1.293	31.600	1.918	62.780	9.876

Figure 5: Performances on Test 1

P	$Ln.1^{st}$	$\sigma(Ln.1^{st})$	$Moves1$	$\sigma(Moves1)$	$Total$	$\sigma(Total)$
5	38.313	1.432	50.920	19.573	84.060	21.018
10	37.077	0.823	49.000	21.832	81.940	23.298
20	36.593	0.664	44.460	14.147	75.680	13.945

Figure 6: Per formances on Test 2

computational effort. This is due to the fact that each time R, T, or "+" moves every individual in both populations must be updated. This means that $P = 20$ performs four times the computational effort than $P = 5$. In Figure 6, one can see once again that $P = 20$ has the best performance. One interesting observation is that the differences in the performances of the three hybrids on Test Environments 1 and 2 with respect to $\overline{Total}$ have remained virtually the same. It will be interesting, in future efforts, to see if this constant remains over a larger test suite of problems.

6 Future Work

At present, we are particularly interested in (1) improving the search efficiency of small population evolutionary motion planners and (2) reducing the computational effort of larger population evolutionary motion planners. We also experimenting with a number of specially designed order-based and relational crossover operators [1, 8] that make better use of knowledge gleaned from the environment. Because of the speed at which these GEPOA-II hybrids develop feasible paths, our future work will be devoted to the development of hybrid motion planning systems that make use of VBR as well as other traditional motion planning concepts. This will allow for the development of evolutionary hybrids that incorporate e the best of both traditional and evolutionary motion planning concepts.

Acknowledgment

This research is partially funded by grants from the NASA Autonomous Control Engineering Center under grant number ACE-48146. The authors wish to thank them for their financial support.

References

[1] Brown. Edward, Jr. (1996). Relational Crossover with Seed Displacement, Technical *Report ACETR-MP-96-03, NASA ACE Center, North* Carolina A *&T State University, Greensboro.*

[2] Ashiru, 1. and Czarnecki, C. (1995). Optimal Motion Planning for Mobile Robots Using Genetic Algorithms, *Proceedings of the 1995 IEEE/IAS International Conference on Industrial Automation and Control.*

[3] Chang, T.-Y., Kuo, S-W., and Hsu, J. Y.-J. (1994). A Two-Phase Navigation System for Mobile Robots in Dynamic Environments, *Proceedings of the* 199.4 *IEEE/RSJ/GI* International *Conference* on *Intelligent Robots and Systems,* pp. 297-300.

[4] Dozier, G., Esterline, A., Homaifar, A., and Bikdash, M. (1996). Hybrid Evolutionary Motion Planning via Visibility-Based Repair, submitted to: *The 1997 IEEE* International Conference on *Evolutionary Computation.*

[q Fujimura, K. (1991). Motion Planning *in* Dynamic *Environments,* Springer-Verlag.

[6] Hocaoglu, C. and Sanderson, A. (1996). Planning Multi-Paths Using Speciation in Genetic Algorithms, *Proceedings of the* 1996 *IEEE* International Conference on Evolutionary *Computation,* pp. 378-383.

[7] Latombe, J.C. (1991). *Robot Motion* Planning, Kluwer Academic Publishers.

[8] McCullough, Shaun. (1996). An Insertion Based PMX Operator for a Genetic Tracking System with Multiple Targets, Technical *Report A CETR-MP-96-04, NASA ACE Center, North Carolina A &T State University, Greensboro.*

[9] Michalewicz, Z. (1994). Genetic *Algorithms + Data Structures = Evolution Programs,* Springer-Verlag.

[10] Radcliffe, N.J. (1991). Forma Analysis and Random Respectful Recombination, *Proceedings of the Fourth International Conference on Genetic Algorithms.*

[11] Shibata, T., and Fukuda, T. (1993). Robot Motion Planning by Genetic Algorithm with Fuzzy Critic, *The 8th IEEE International Symposium on Intelligent Control,* pp. 565-570.

Lateral forces between a scanned probe and surface

39/25

Mary A. Drummond Roby (National Science Foundation Graduate Fellow)
and G. C. Wetsel
Nanoelectronics Laboratory
The University of Texas at Dallas

The interaction between a scanned probe and a sample can be utilized to obtain images of surface topography and investigate forces on a submicron scale. If the scanned probe is a near-field optical microscope the lateral force between the tip and surface serves two purposes: it provides a mechanism upon which to regulate the tip-sample separation and also provides information on the surface topography. This provided a "bonus" topographical image to accompany the optical image. More recently, the lateral vibration of the of the probe has been used for force imaging since it does not suffer from the same "snap" to the sample's surface as AFM's which are oriented in the same plane as the surface. Avoiding the "snap" also allows an NSOM to be used to investigate samples submersed in a fluid[2] or obtain high-lateral-definition images of liquid-crystal films[3].

We have previously reported[1] the development of a technique for calibrated determination of the lateral force on a vibrating probe. A dynamic continuum model was used in lieu of a harmonic oscillator to take advantage of the physical boundary conditions, including the intrinsic loss of the fiber vibrating in a fluid and a friction force acting on the tip. The frictional force has been experimentally evaluated -- in the far field, the force is due to the fluid trapped between the probe and surface'and, in the near-field, the force is elastic in nature[6]. In this paper we describe the modeling of the probe and the nature of the forces at the probe's tip. Additionally, preliminary data is presented indicating that the forces are dependent on surface species.

The computer controlled experiment includes a single-mode silica fiber mounted to a segmented piezoelectric (lead-zirconium-titanate) tube[7] such that the fiber's motion is parallel to the surface. The tube is driven by a function generator. A HeNe laser is focused onto the fiber near its tip. On the opposing side of the fiber, a position-sensing-detector (PSD) monitors the displacement of the scattered light by the fiber. The signal ii-cm the PSD is amplified, and the in-phase and phase-quadrature signals from the PSD are measured using a lock-in amplifier and recorded by the computer. The PSD's voltage output is calibrated in terms of fiber displacement by moving the non-vibrating fiber a known distance and recording the PSD's voltage output. Thus, the amplitude of the PSD signal is converted to fiber displacement using the PSD displacement calibration. The sample's coarse approach is controlled with a burleigh pusher[8] or picomotor[9]. The fine approach is controlled by a second segmented transducer. As the sample is approached, the fiber is vibrated at its resonant frequency, and the in-phase and phase-quadrature are monitored as described above. The force at the end of the probe is determined from this data

The dynamical model is a continuous elastic rod vibrating in flexure (bending). The transverse displacement (u) from equilibrium is assumed to be small compared to the length (L) of the rod, and the wavelength of the vibration is assumed to be huge compared to the diameter (2a) of the rod. Intrinsic losses--due to internal attenuation and (principally) the viscous drag of air--are included knowledge of the intrinsic loss is important to determination of the surface interaction and to calculation of the sensitivity of the method The equation of motion is[10,11]:

$$\frac{\partial^2 u}{\partial t^2} + v_l^2 K_0^2 \frac{\partial^4}{\partial z^4}\left(u + \alpha \frac{\partial u}{\partial t}\right) = 0, \tag{1}$$

where $v_l=(E/\rho)^{1/2}$ is the longitudinal-wave velocity, $K_0=a/2$ is the radius of gyration, E is Young's modulus, p is the density, and a is the intrinsic-loss factor. For harmonic time dependence, all dynamical quantities vary with time as $e^{j\omega t}$; the spatial solution is:

$$u(z) = A\cos\gamma z + B\sin\gamma z + C\cosh\gamma z + D\sinh\gamma z. \tag{2}$$

where

$$\gamma^4 = \frac{\omega^2}{v_I^2 K_0^2 (1 + j\omega\alpha)} \tag{3}$$

Solutions to (2) are of the same form[10,11] as for $\alpha=0$ except for the complex argument corresponding to (3):

The boundary conditions are used to determine the 4 arbitrary constants in Eq. (2). The displacement is specified at z=0: $u(0)=u_0$. The angular displacement, $\partial u/\partial z$, is assumed to be zero at z=0. The torque, $EI(\partial^2 u/\partial z^2)$, is assumed to be zero at z=L, where I is the moment of area about they axis; hence, $(\partial^2 u/\partial z^2)_{z=L}=0$. The third derivative of u is related to the lateral force acting at z=L: $(\partial^3 u/\partial z^3)_{z=L}=-F(L)/EI$. Thus, the displacement can be related to the exciting displacement u_0, the frequency ω, material parameters, and the lateral force F(L). For the data analyzed, it was assumed that the lateral force is a frictional force proportional to the particle velocity at the end of the fiber: $F(L)=-\eta(\partial u/\partial t)_{z=L}$, where η is a frictional-force coefficient.

In order to evaluate the model and method of dynamic lateral force determination, frequency responses of the fundamental mode of vibration of silica single-mode optical fibers[12] with the cladding removed were measured. The results are illustrated by the data in Fig. 1. The set of discrete points in Fig. 1(a) represents the measured frequency response of the fiber before interaction with the surface. The solid curve plotted through those points represents the theoretical fit of Eq. (2) with $\eta=0$; the values of a and L were varied to obtain the minimal RMS deviation of the fit and the data. The measured frequency response when the fiber interacted with the surface is represented by the set of discrete data points in Fig. 1(b). The solid curve through the experimental data represents the minimal-RMS-deviation fit of Eq. (2); the values of a and L determined above were used and η was varied for a best fit.

Since a significant effect on the fiber's motion is observed in air when the end of the fiber is several μm's from the sample surface, the distance between fiber tip and surface is small compared to the diameter of the fiber. With this geometry, a one-dimensional model such as that shown in Fig. 2 can be used. The oscillating plate at x=0 is separated from a fixed plate at x=h by two layers of viscous fluid characterized by densities, ρ_1 and ρ_2, and by kinematic

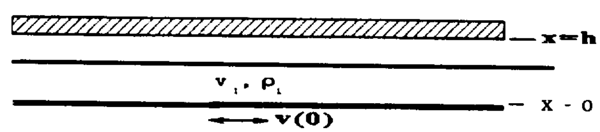


Figure 2

viscosities, v_1 and v_2. The moving plate represents the tip and the layer near it is taken to be air. The fixed plate represents the sample and the layer near it is taken to be contamination, such as water. The equation of motion for an incompressible viscous fluid is given by the Navier-Stokes equation[13]

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} = \frac{-1}{\rho}\nabla p + v\nabla^2\boldsymbol{v}, \tag{4}$$

where v is the fluid particle velocity and p is the fluid pressure. For this geometry, $\boldsymbol{v}(x,t) = \hat{j}v(x)e^{i\omega t}$ where the particle motion is assumed to be in the y direction (parallel to the plates) and ω is the angular frequency of vibration of the fiber. Substitution of v(x,t) into (4) reveals that the pressure is uniform, $\partial p/\partial x=0$. The solutions to (4) for each layer are

$$v_j(x) = A_j e^{-K_j x} + B_j e^{+K_j x}, \tag{5}$$

where $K_j=(1+i)/\delta_j$, and $\delta_j=(2v_j/\omega)^{1/2}$, j= 1,2. These solutions represent evanescent waves and the δ_j are the e^{-1} penetration

depths.

The constants A_1, B_1, A_2, and B_2 are evaluated using the boundary conditions. There is no relative motion at the interface between a viscous fluid and a solid, therefore, $v_2(h)=0$, and $v_1(0)=v_0$, where v_0 is the particle velocity of the oscillating plate. The velocity is continuous at the interface between the two fluids: $v_2(d)= v_1(d)$. Furthermore, the stress is continuous at the fluid interface. Determination of the constants using the boundary conditions allows the fictional force on the oscillating plate to be calculated[13]:

$$F(0) - \frac{\nu_1 S_f}{\rho_1}\left(\frac{\partial v_1}{\partial x}\right)_{x=0}, \tag{6}$$

where S is the area of the plate. By combining the two equations for F(0), the fictional force coefficient can be expressed as a function of the parameters of the two-layer model of the fluids between the tip and the sample, $\eta=-F(0)/v_0$. Figure 3 shows the experimentally determined η compared with the theoretical η determined from Eq. (6) using the area of the fiber and the physical parameters for air as the fluid between the two plates.

Once fluid dynamics were determined for the damping effect on the probe's frequency response, additional force spectroscopy with a smaller diameter probe were performed. At closer distance, the frequency response of the probe was not only damped but there was a notable upward shift in the resonant frequency. Since the fluid dynamic model does not account for a frequency shift, anew force must be responsible. By assuming that the torque at z=L was zero, the lateral force is modified to include an elastic restoring force proportional to displacement.

$$F(L)= -\eta\left(\frac{\partial u}{\partial t}\right)_{z=L} \kappa u(L), \tag{7}$$

The phenomenological parameter, κ, **might** represent the elastic deformation of the surface and/or surface-tension of a thin liquid layer on the surface.

The new lateral force model was evaluated by measuring the frequency response of the fiber between translations of the sample. The amplitude of the measured fiber displacement is shown in Fig. 4 as a function of frequency for representative values of sample extension (Ah). For each value of Ah, the experimental data is compared with a theoretical fit of Eq. (7) by varying η and κ. Curves (a), (b), and (c) represent probe-sample separations corresponding to the far field. For these distances, the contribution of the elastic force is negligible and the viscous force is sufficient to account for the force at the tip; the viscous force produces no change in resonant frequency under these conditions Notice that the experimental resonant frequency is the same for these values of Ah. Curves (d), (e), and (f) represent smaller values of probe-sample separation, notice that there is now an increase in the resonant frequency. For these values of Ah, the contribution of the elastic force is not negligible; the viscous force is insufficient to account for the force at the tip. It is the elastic force that causes the change in resonant frequency. Figure 5 shows the distance dependence for both the fluid, elastic, and total force.

In determining the nature of the force at the end of the vibrating fiber, the sample was a piece of silicon wafer. However, other materials such as photoresist and quartz were "investigated. The preliminary approach curves shows that the force depends on the material. In Figure 6, the normalized magnitude of approach curves on quartz and silicon are similar, but the curve for photoresist is more abrupt. The phase of the force also varies as show in Figure 7. On silicon, the phase change is approximately 100 while the phase changes by 40° on silicon. These differences in both magnitude and phase indicated that either the fluid or elastic force maybe linked to the surface species of the sample.

1. E. Betzig, P.L. Finn, and J.S. Weiner, "Combined shear force and near-field scanning optical microscopy", *Appl. Phys. Lett. 60, 2484-2486(1 992).*
2. Patrick J. Moyer and Stefan B. Kämer, "High-resolution imaging using near-field scanning optical microscopy and shear force feedback in water", *Appl. Phys. Lett.*, 68, 3380-3382 (1996).
3. R.L. Williamson and M.J. Miles, "Studying topography and subsurface structure in 8CB liquid crystal films with shear-force microscopy", *J. Appl. Phys.* **80**, *3143-3146 (1996).*
4. G. C. Wetsel, Jr. and M.A. Drummond Roby, "Dynamic nanoscale lateral-force determination", *Appl. Phys. Lett. 67, 2735-2737 (1995).*
5. M.A. Drummond Roby, G.C. Wetsel, Jr., and C.-Y. Wang, "Scanned-probe lateral-force determination of fluid-dynamic effects near a solid surface in air", *Appl. Phys. Lett.* 69,130-132 (1996).
6. M.A. Drummond Roby and G.C. Wetsel, Jr., "'Measurement of elastic force on a scanned probe near a solid surface",

7. G. Binnig and D.P.E. Smith, "Single-tube three-dimensional scanner for scanning tunneling microscopy", Rev. *Sci. Instrum. 57, 1688-1689 (1986).*
8. Burleigh Instruments, Inc., Burleigh Park, Fishers, NY 14453.
9. New Focus, inc., Santa Clara, CA 95051.
10. J. W. Strutt (Baron Rayleigh), *The Theory of Sound*, Vol. J, Dover, New York (1945).
11. L.D. Landau and E.M. Lifshitz, *Theory of Elasticity*, 3'rd Ed., Pergamon Press, oxford (1986)
12. Corning Incorporated, Corning, NY 14831.
13, L.D. Landau and E.M. Lifshitz, *Fluid Mechanics*, Pergamon Press, Oxford (1959).

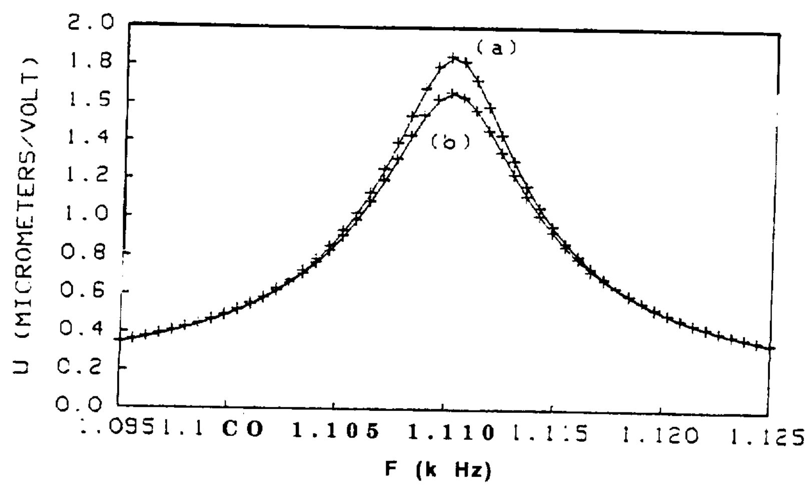


Figure 1. The discrete points represent the measured frequency response of a quartz fiber (a) in air, and (b) in contact with a quartz sample. The solid curves represent the theoretical fits of Eq.(4) using L=9.5365 mm, a=62.5 μm. The value of the intrinsic damping factor inferred from the fit to the data (a) is a =7.345x10 '’s-'. The value of the frictional-force damping factor inferred from the tit to the data of(b) is η=2.6x1 O '’N-s/m.

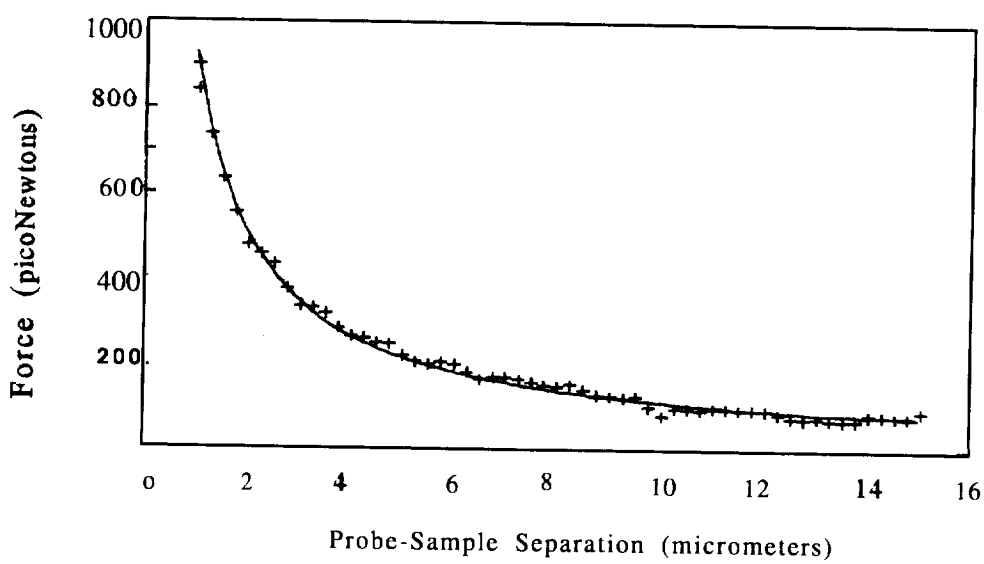


Figure 3. Force on the fiber tip **vs** probe-sample separation (A *h*): experimental (plus) and theoretical (solid).

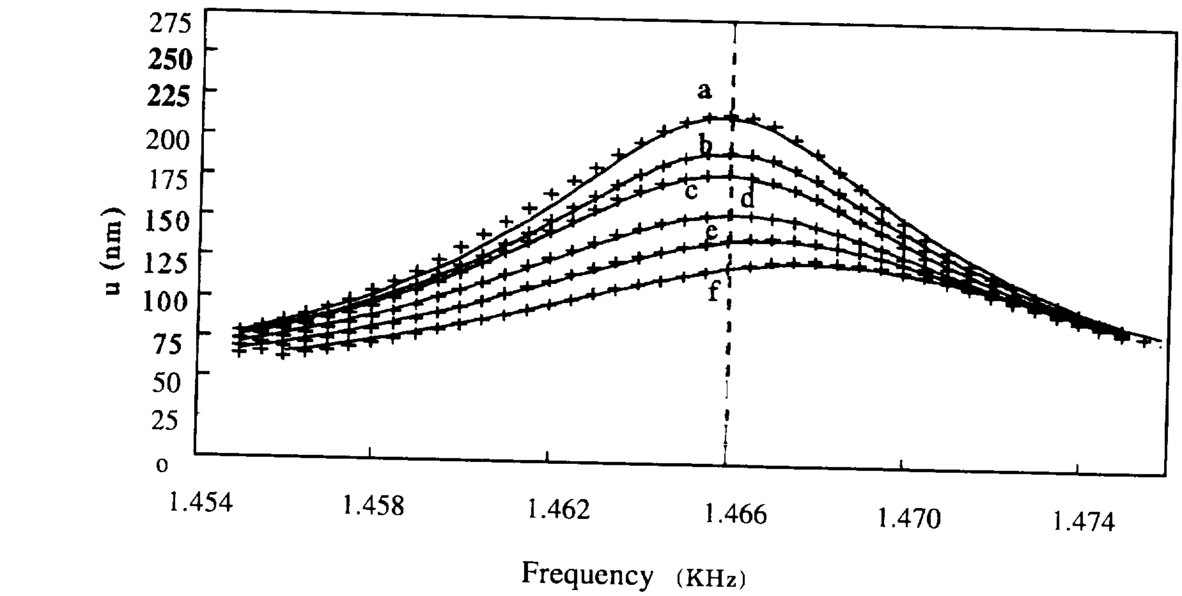


Figure 4. Frequency response of displacement of fiber for several values of sample extension, A h. The discrete points represent the data and the solid curve represents a minimal-RMS-deviation fit of Eq. (2).

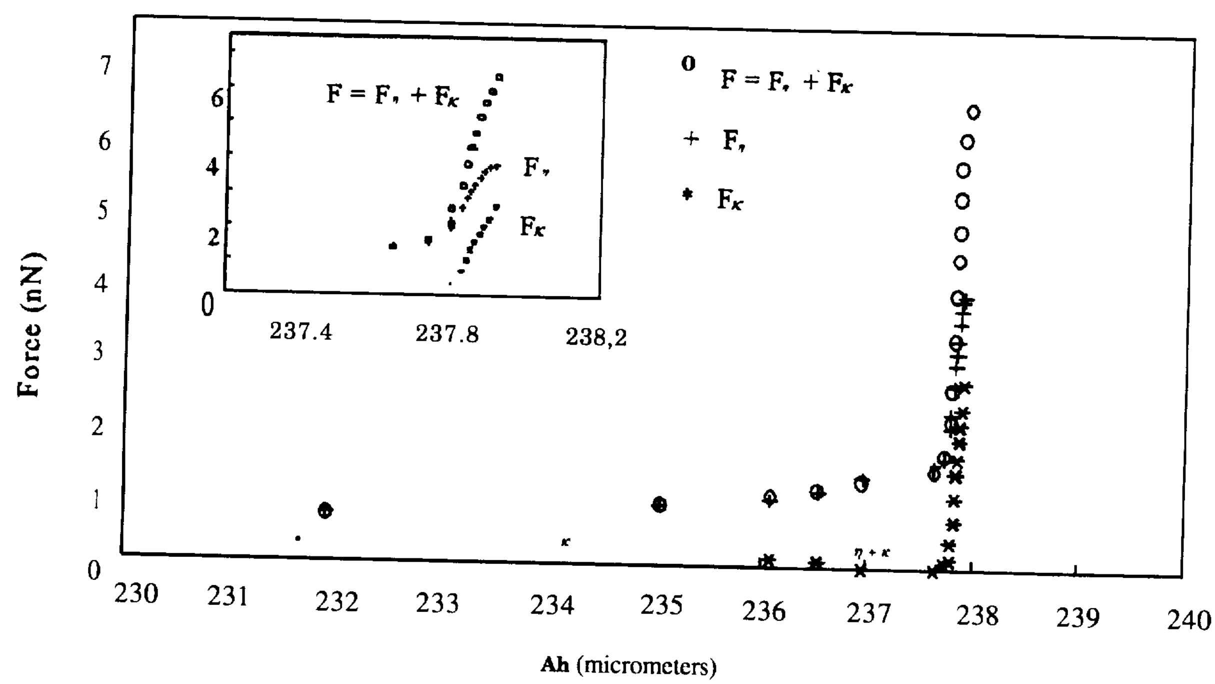


Figure 5. Viscous force, F_η (L), elastic force, F_ε (L), and total force, $F_{\eta+\varepsilon}$ (L) vs Ah. Insert is a close-up of transition region.

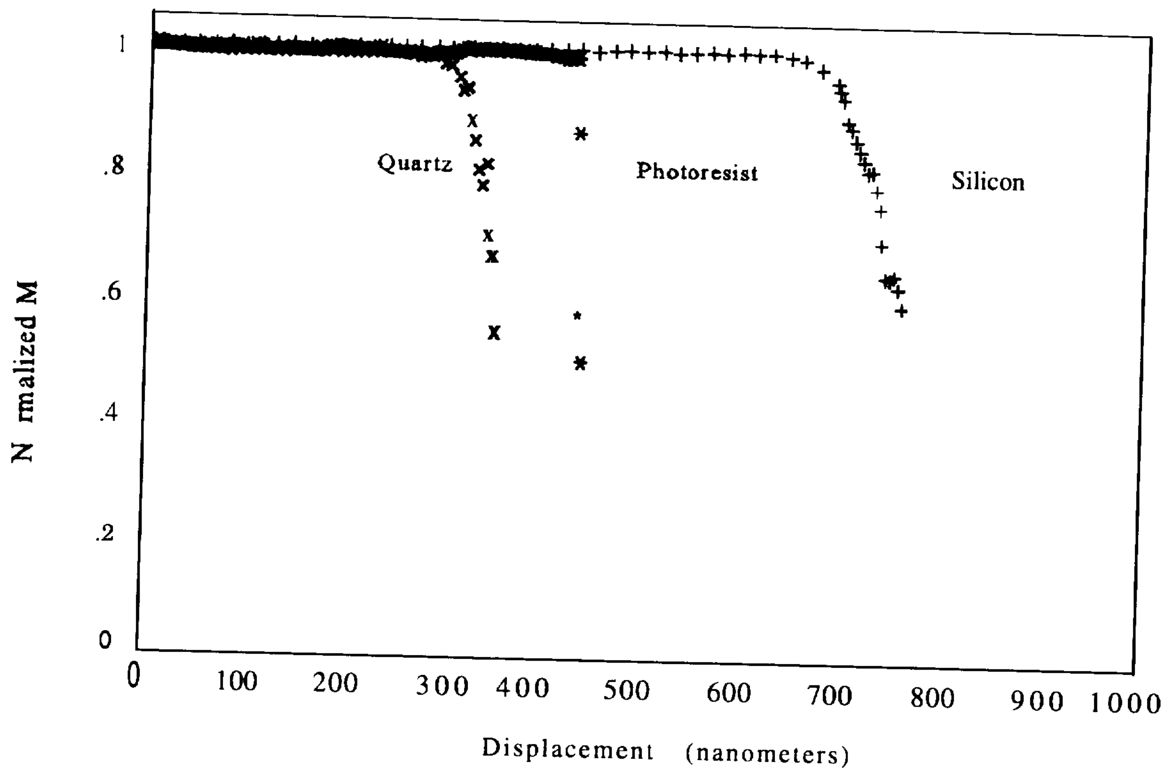


Figure 6 . Normalized magnitude versus displacement for quartz, **photoresist**, and silicon.

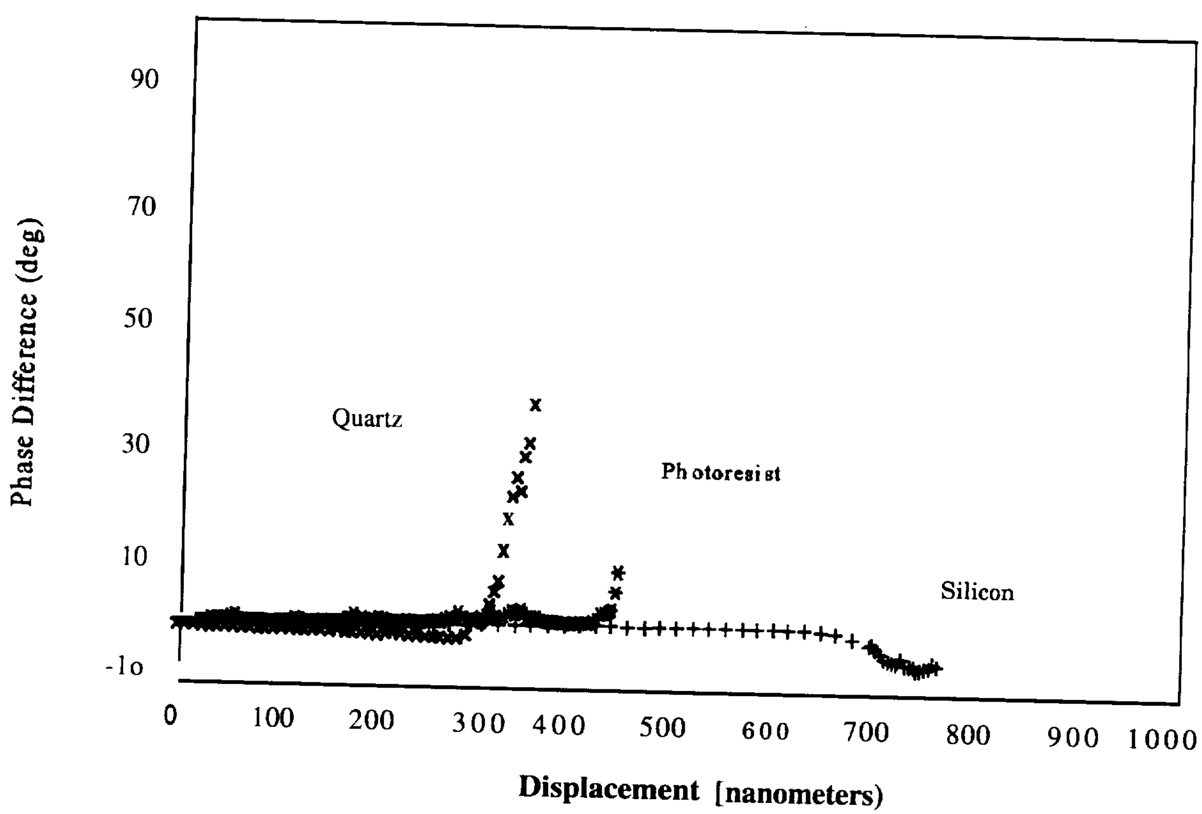


Figure 7 . Phase difference versus displacement for quartz, **photoresist**, and silicon.

URC97040

Fuzzy Spatial Reasoning

A.C. Esterline G. Dozier A. Homaifar

NASA Center for Autonomous Control Engineering
North Carolina A&T State University
Greensboro, NC 27411, USA
esterlin@ncat.edu

Abstract

We present a fuzzy version of the crisp spatial logic developed by Randell et al., which takes the single relation connected-with *as primitive. Membership functions are defined for each spatial relation defined in the crisp theory. Furthermore, principles are presented for defining linguistic variables whose linguistic values are spatial relations. The work reported here addresses spatial reasoning in situations where numerical or geometric precision is unlikely; it is particularly suited for dynamic situations.*

Keywords: Spatial Reasoning, Fuzzy Logic, Ontology, Mereology, Topology, Linguistic Variables, Lattices, Constraints, Qualitative Simulation

1. Introduction

Spatial Reasoning is central to several important areas of engineering and computer science, including computer vision, assembly, and especially path planning for mobile robots. Because of the sort of mathematics forming the ubiquitous background of engineers, it is natural that engineers should attempt to represent spatial relations in two- or three-dimensional Euclidean space with values from the real number system or vector spaces over the reals or, in certain domains, with geometric structures. The data available in many situations, however, are not sufficiently complete or precise to support numerical or geometric descriptions that truly reflect the actual state of affairs. Furthermore, if human observers or controllers are involved, the data are likely in a linguistic form that is only indirectly related to numerical or geometric representations. The route of escape from this unnecessary and often spurious tyranny of precision without abandoning rigor is in the direction of other, often more modem, areas of mathematics, such as topology, abstract algebra (especially lattice theory and Boolean algebras), and formal logic. Although these areas avoid imposing notions of precision that are out of place, they reveal, often more perspicuously than does traditional engineering mathematics, relations of consistency, inconsistency, and entailment. Further insight into spatial relations has arisen from ontology, the study of the basic categories of what exists, which is traditionally an area of philosophy [7] and has recently been extensively applied in artificial intelligence [6]. Within ontology, we are particularly concerned with the areas called mereology, which addresses the par-of (or part-whole) relation, and "topology", which addresses the connected-with relation and, as the term is used here, uses different methods to address roughly the same area as that addressed by algebraic topology in mathematics [8]. Following Randell, Cohn, and Cui [2,7], we shall take the connected-with relation as primitive and derive from it an entire system of spatial relations, including the part-of relation. The entire system is expressed in first-order predicate logic and includes special axioms from which the fundamental properties of interest are derived. This crisp theory involves no numbers and does not directly support geometric properties. Yet it allows one to reason about spatial relations and what sequences of these relations are possible as a system evolves.

The crisp theory of spatial relations, however, is not entirely free from dependence on precision, for many of the relations occur at thresholds and all of them may shade imperceptibly into other relations. This dependence on precision becomes onerous when the theory is applied to complex systems, about which we have limited ability to make precise statements if only because we usually begin with imprecise data. This motivates introduction of a fuzzy version of the Randell *et al.'s* spatial logic since fuzzy set theory and logic adapt crisp notions from set theory and logic so that they may accommodate imprecision and approximation [9]. We do this, however, in a principled way, so as not to compromise rigor. The general principle followed here is to develop fuzzy logics from already proven crisp logics in a way that preserves their desirable properties.

Much of the appeal of a logical formalism is its close correspondence with natural language. This is preserved with the fuzzy version. Indeed, this correspondence is enhanced by the linguistic variables of the fuzzy version as long as the values are motivated by the logical structures themselves and are not derived solely from convenience in representing membership functions.

The remainder of this paper is organized as follows. In section 2, we present the basics definitions of Randell *et al.*'s *crisp theory* and the definitions of the membership functions for the fuzzy version. Section 3 shows how linguistic variables should be defined for the fuzzy theory. Section 4 briefly addresses the methods of reasoning that are hereby enabled, and section 5 concludes and addresses future work.

2. The Basic Definitions

We assume one dyadic (two-argument) relation, $C(x,y)$, understood to mean that x is connected with y. This relation is taken to include any kind of spatial contiguity, include the cases where x is a part of y, where x and y overlap, and where the boundaries of x and y share a common point but x and y have disjoint interiors. In terms of point-set topology, $C(x,y)$ holds when the closures of regions x and y share a common point. The C relation and the other relations to follow are sufficiently abstract that the individuals x and y can be thought of as several kinds of things, for example, physical objects, regions in space, and even temporal regions. To form examples of theses relations, it is helpful to think of the individuals as either spatial regions with physical boundaries (such as rooms) or objects that can be located in such regions (such as desks in rooms).

Two properties of the connected-with relation are immediately evident. First of all, it is reflexive: anything is connected with itself, or, in symbols, $\forall x\, C(x,x)$. Secondly, the connected-with relation is symmetric: if x is connected with y, then y is connected with x, or, in symbols, $\forall xy\, [C(x,y) \rightarrow C(y,x)]$.

The translation of these properties to membership fictions for dyadic fuzzy relations is straightforward. We assume a universe U of individuals to which our spatial relations apply. Then, where $U\times U$ is the Cartesian product of U *with itself* (that is, the set of all ordered pairs where both the first and the second elements are in U), any dyadic fuzzy relation R on U *has the* form

$$R = \{((x,y), \mu_R(x,y)) | (x,y) \in U \times U\}$$

That is, R *is viewed as the* set of pairs (x,y), with $x \in U$ and $y \in U$, each pair itself paired with a value of the membership function μ_R for R. We *assume* that every fuzzy relation R *we* work with is normal, that is, $0 \leq \mu_R(x,y) \leq 1$ for all $x,y \in U$. We require that the crisp version of the spatial logic be recoverable from the fuzzy version. To that end, we define the crisp relation corresponding to the fuzzy relation R *as the* strong 0.5–cut $R_{0.5}$ of R:

$$R_{0.5} = \{(x,y) \in U \times U | \mu_R(x,y) > 0.5\}$$

That is, the relation $R_{0.5}$ holds from x to y (or between x and y, in that order) if and only if $\mu_R(x,y) > 0.5$. Note that $R_{0.5}$ *is* then a subset of $U\times U$, *as* required of a crisp relation over U. Translating to the crisp version, then, amounts to taking a formula of the form $R'(a,b)$ (*a is* related by R 'to b – for example, *a is* connected with b) and considering it true if $\mu_R(a,b) > 0.5$, where R *is the fuzzy* relation corresponding to R ! *If* $\mu_R(a,b) < 0.5$, then $R'(a,b)$ *is* considered false. Thus, if we begin with a crisp version in which all statements of the form $R'(x,y)$ *are* true, *we* translate to the fuzzy version by considering a relation

$$R = \{((x, y), \mu_R(x,y)) | (x,y) \in U \times U\}$$

where we require $\mu_R(x,y) > 0.5$ if $R'(x,y)$ *is true* (or $(x,y) \in R'$). Thus, the natural way to express the reflexive property of the fuzzy connected-with relation is

$$\mu_C(x,x) > 0.5$$

where, as throughout, we use the same name for the fuzzy relation as for the crisp relation - context clearly disambiguates.

Next, the natural way to view reflexivity is as a guarantee that the order of the arguments makes no difference. Thus, for the fuzzy connected-with relation, we require that

$$\mu_C(x,y) = \mu_C(y,x)$$

Given the connected-with relation $C(x,y)$, we can define a basic set of dyadic spatial relations. First of all, we say that x is disconnected from y (written $DC(x,y)$) if x is not connected with y:

$$DC(x,y) =_{def} \neg C(x,y)$$

Since we require our fuzzy relations to be normal, the appropriate notion of negation is complement with respect to 1. Thus we require of the fuzzy DC relation that

$$\mu_{DC}(x,y) = 1 - \mu_C(x,y)$$

Next, x is a part of y, $P(x,y)$, if everything connected with x is also connected with y:

$P(x,y) =_{def} \forall z\ [C(z,x) \rightarrow C(z,y)]$

The corresponding constraint on the fuzzy relation P can be developed in a small number of steps. First, we think of a conditional $\varphi \rightarrow \psi$ in the manner of material implication as $\neg\varphi \vee \psi$. As is usual, we relate the *OR*, v, with the maximum of the membership functions. Similarly, we relate the *AND*, $\wedge$, with the minimum of the membership functions. A universal quantification $\forall x\ \varphi(x)$ is thought of as the conjunction of all substitution instances of $\varphi(x)$: $\varphi(u_1) \wedge \varphi(u_2) \wedge \ldots \wedge \varphi(u_n)$, where $U = \{u_1, u_2, \ldots, u_n\}$. Thus, universal quantification is related with the minimum value of the argument as the variable bound by the quantifier ranges over the elements of the universe. Similarly, an existential quantification $\exists x\ \varphi(x)$ is thought of as the disjunction of all substitution instances of $\varphi(x)$: $\varphi(u_1) \vee \varphi(u_2) \vee \ldots \vee \varphi(u_n)$. Thus, existential quantification is related with the maximum value of the argument as the bound variable ranges over the elements of the universe. Following these principles, we require for the fuzzy relation P that

$$\mu_P(x,y) = \min_{z\in U}\left\{\max\left\{1-\mu_C(z,x), \mu_C(z,y)\right\}\right\}$$

The definitions of the remaining crisp relations can be related to constraints on the membership functions of the corresponding fuzzy relations following the principles articulated above. We now list these relations without further comment. For each relation, we first give the English reading, then the definition of the crisp relation, and finally the constraint on the corresponding fuzzy relation. Figure 1 portrays the more specific of the spatial relations.

x is a proper part of y

$PP(x,y) =_{def} P(x,y) \wedge \neg P(y,x)$

$$\mu_{PP}(x,y) = \min\left\{\mu_P(x,y), 1-\mu_P(y,x)\right\}$$

x is identical with y

$x = y =_{def} P(x,y) \wedge P(y,x)$

$$\mu_=(x,y) = \min\left\{\mu_P(x,y), \mu_P(y,x)\right\}$$

x overlaps y

$O(x,y) =_{def} \exists z\ [P(z,x) \wedge P(z,y)]$

$$\mu_O(x,y) = \max_{z\in U}\left\{\min\left\{\mu_P(z,x), \mu_P(z,y)\right\}\right\}$$

x partially overlaps y

$PO(x,y) =_{def} O(x,y) \wedge \neg P(x,y) \wedge \neg P(y,x)$

$$\mu_{PO}(x,y) = \min\left\{\mu_O(x,y), 1-\mu_P(x,y), 1-\mu_P(y,x)\right\}$$

x is discrete from y

$DR(x,y) =_{def} \neg O(x,y)$

$$\mu_{DR}(x,y) = 1-\mu_O(x,y)$$

x is externally connected with y

$EC(x,y) =_{def} C(x,y) \wedge \neg O(x,y)$

$$\mu_{EC}(x,y) = \min\left\{\mu_C(x,y), 1-\mu_O(x,y)\right\}$$

x is a tangential proper part of y

$TPP(x,y) =_{def} PP(x,y) \wedge \exists z\ [EC(z,x) \wedge EC(z,y)]$

$$\mu_{TPP}(x,y) = \min\left\{\mu_{PP}(x,y), \max_{z\in U}\left\{\min\left\{\mu_{EC}(z,x), \mu_{EC}(z,y)\right\}\right\}\right\}$$

x is a nontangential proper part of y

$NTPP(x,y) =_{def} PP(x,y) \wedge \neg\exists z\ [EC(z,x) \wedge EC(z,y)]$

$$\mu_{NTPP}(x,y) = \min\left\{\mu_{PP}(x,y), 1 - \max_{z \in U}\left\{\min\left\{\mu_{EC}(z,x), \mu_{EC}(z,y)\right\}\right\}\right\}$$

The relations *P, PP, TPP,* and *NTPP* are non-symmetric. For example, if x is a part of y (that is, $P(x,y)$), it does not follow that y is a part of x. Thus, the inverse of any of these relations is distinct from the relation itself. For example, asserting that x is a part of y asserts something different from asserting that y is a part of x. Where Φ is any relation, we represent its inverse by Φ^{-1}. Concerning the constraints on the corresponding fuzzy relations, note that, in general,

$$\mu_{\Phi^{-1}}(x,y) = \mu_{\Phi}(y,x)$$

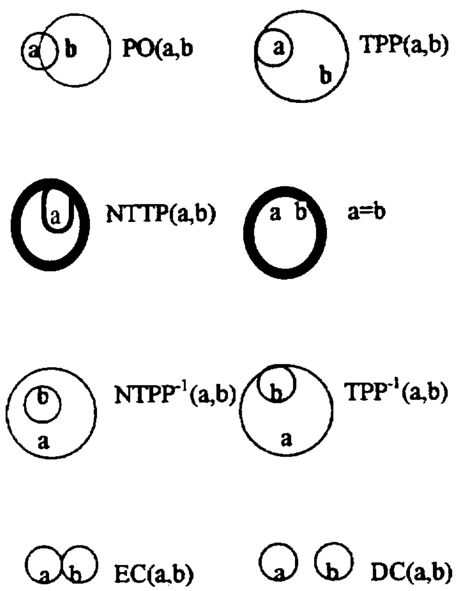


Figure 1. Some spatial relations

3. *Linguistic Variables*

As pointed out by Randell *et* al. [7], the entailment relations among the spatial relations induce a lattice structure on the family of spatial relations. We exploit this lattice to define spatial linguistic variables. Figure 2 shows the lattice of spatial relations. Where R_1 and R_2 *are* spatial relations, we write $R_1 \leq_L R_2$ if R_1 *is* below R_2 in the lattice, that is, if there is a path from R_1 to R_2 going in the direction of the top, $\top$. Note that, for any relation R, *we have* $R \leq_L R$, that is, $\leq_L$ is reflexive. For any relations R_1 and R_2 such that $R_1 \leq_L R_2$, *we have*

$$\forall xy\ [R_1(x,y) \rightarrow R_2(x,y)]$$

The top relation in the lattice, $\top$, is the tautologous relation, which holds of any pair of elements: $\forall xy\ \top(x,y)$. The bottom relation, $\bot$, is the contradictory relation, which holds of no pair of elements: $\forall xy\ \neg\bot(x,y)$.

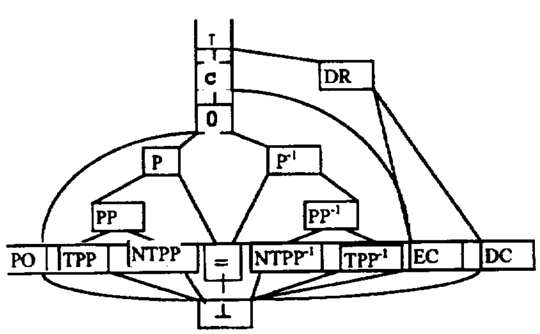


Figure 2. The lattice of spatial relations [7].

Now, let us define a *cutset* of the lattice to be a set of relations that has as a member one relation from every path between $\perp$ and $\top$. Examples of cutsets are *{C, DR}*, *{PO, P, P^{-1}, EC, DC}*, and *{PO, TPP, NTPP, =, NTPP^{-1}, TPP^{-1}, EC, DC}*. The latter set is of particular note since it comprises the relations immediately above $\perp$; we denote it by B. It can be shown that the members of any cutset are exhaustive and pairwise disjoint [7]. That is, given a pair x, y of individuals, exactly one relation in a cutset holds of that pair.

Turning to the corresponding fuzzy relations, the exhaustive and exclusive nature of the relations in a cutset is reflected in the fact that, given any pair x, y of individuals, the value $\mu_R(x,y)$ of the member function of one relation R for x and y tends to clearly dominate the values for x and y of the member functions of the other relations in the cutset as long as the crisp condition is approached. As imprecision and uncertainty take over, of course, this dominance disappears and there maybe several contenders for the dominant relation. The characteristics of this dominance motivate defining a linguistic variable corresponding to $\top$ whose linguistic values are the relations in a given cutset. Such a linguistic variable acts as a classifier; the cutset selected for the linguistic values determines how finely and in what respects we classify. In particular, if B is used, we have the finest classifier supported by this taxonomy of spatial relations.

We can define linguistic variables that are restricted to a part of the lattice by introducing the notion of an *R-cutset*, where R is a spatial relation. An R-cutset is simply a cutset of the sublattice consisting of those relations R_i such that $R_i \leq_L R_j$; note that R is the top of the sublattice. For example, there are two P-cutsets: *{PP, =}* and *{TPP, NTPP, =}*. Given a relation R, we can define a linguistic variable whose linguistic values are the relations in a given R-cutset. Such a linguistic variable acts as a restricted classifier.

Use of linguistic variables in a principled way, such as here, offers several advantages. For one thing, linguistic variables greatly facilitate human input and allow output that is easily interpreted by humans. Again, linguistic variables support discrete classification allowing a *system* to focus on the most significant aspects. Using restricted linguistic variables, based on R-cutsets, further supports the ability of a system to focus on significant aspects. Finally, allowing different cutsets or R-cutsets to supply linguistic values supports flexibility while retaining access to all relations in the taxonomy.

4. Reasoning with the Fuzzy Spatial Logic

Randell *et al.* support reasoning with the crisp spatial logic by using a theorem prover for many-sorted logic [7]. Our approach with the fuzzy version, in contrast, exploits the connections among the membership functions. These connections form a network of constraints for which we are implementing constraint satisfaction methods.

Randell *et al.* have used their spatial logic for qualitative simulation of dynamic situations [6]. Figure 3 shows the allowable transitions among the relations of cutset B; only these transitions should occur as two bodies move or otherwise change in relation to each other. The same transitions are allowable for the fuzzy relations, but now precise thresholds need not be enforced. We are particularly interested in applications to dynamic path planning for mobile robots.

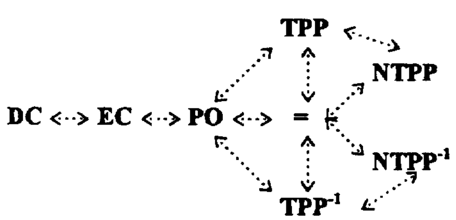


Figure 3. Transitions among relations

5. Conclusions and Future Work

We have presented a fuzzy version of the crisp spatial logic developed by Randell *et al.* that takes the single relation connected-with as primitive. Membership functions have been defined for each spatial relation defined in the crisp theory. Furthermore, principles have been presented for defining linguistic variables whose linguistic values are spatial relations.

To apply our results, we must develop principles for how various information sources – such as sensors and human speech – can contribute to the fuzzy measures. To exploit our results, we are developing constraint

satisfaction methods to apply to the constraint network induced by the definitions of the membership functions. For such methods to be feasible, we need ways to limit the number of individuals considered. This is particularly critical for cases where the crisp definitions include quantifiers, for then we take maxima and minima over all the individuals in our universe of discourse. Intuitively, however, one should be able to exploit known spatial relations to safely ignore the vast majority of individuals in anyone case.

The spatial logic of Randell *et al.* is only one of several AI logics for which useful fuzzy versions can be developed. We intend to address Allen's logic of temporal intervals [1] and Galton's logic of space, time, and motion [3].

Bibliography

[1] Allen, J.F., An interval-based representation of temporal knowledge, *IJCAI* '81,221-226.

[2] Cohn, A. G., Randell, D. A., & Cui, Z., Taxonomies of logically defined qualitative spatial relations, *Int. Journal of Human-Computer Studies, 43 (1995), 831-846.*

[3] Galton, A., Towards an integrated logic of space, time, and motion, IJCAI '93, 1993, 1550-1555.

[4] Guarino, N. and Poli, R (eds.), *The Role of Formal Ontology in the Information Technology, special issue* of the *International Journal of Human-Computer Studies, 43, 5/6 (Dec. 1995).*

[5] Johansson, I., *Ontological Investigations: An Inquiry into the Categories of Nature, Man and Society.* London Routledge, 1989.

[6] Randell, D.A., Cohn, A. G., and Cui, Z., Naive topology: modeling the force pump, in B. Fakings and P. Struss (eds.), *Recent Advances in Qualitative Physics.* Cambridge, MA: MIT Press, 1992, 177-192.

[7] Randell, D. A., Cui, Z., and Cohn, A. G., A spatial logic based on regions and connection, in B. Nebel, C. Rich, and W. Swartout (eds.), *KR* 92 San Mateo, CA: Morgan Kaufmann, 1992, 165-176.

[8] Varzi, A., On the boundary between mereology and topology, in R Casati, B. Smith, and G. White (eds.), *Philosophy and the Cognitive Sciences.* Vienna Hölder-Pichler-Tempsky, 1994.

[9] Zimmermann, H., *Fuzzy Set Theory and its Applications,* 3rd ed. Boston: Kluwer, 1996.

11/80

The HSCaRS Summer Enrichment Program; Research Opportunities for Minority and Women Undergraduates in Global Change Science

Authors: Maurice G. Estes, Jr./IGCRE, Donald J. Perkey/IGCRE, T. L. Coleman/Alabama A&M University

Background and Program Description

The Center for Hydrology, Soil Climatology, and Remote Sensing (HSCaRS) was established July 1, 1995, through a cooperative agreement between the National Aeronautics and Space Administration (NASA) and Alabama A&M University. One challenge in the Center was to develop an educational component that would increase participation by students in mainstream research and increase the production of underrepresented minorities who are U.S. citizens in NASA-related fields. This goal was strongly supported by a number of educational research publications. The Task Force on Women, Minorities, and the Handicapped in Science and Technology reported in the 1988 Interim Report entitled, *Changing America: The New Face of Science and Engineering*, that one of America's most urgent tasks is to strengthen the science and engineering workforce. The Task Force asserts that, "by the year 2000, 85 percent of new entrants to the Nation's work force will be members of minority groups and women". According to the Wall Street Journal, summer internships are flourishing with an emphasis on minorities. Bausch and Lomb Inc., and Chervon Corporation are increasing summer internship opportunities and are focusing recruiting efforts on minority candidates. The Aluminum Company of America is focusing internship opportunities on minorities and women with engineering and technical training. The challenge is clear, more women and minority scientists must be educated to meet the needs of America's technics] work force.

Based on this data, a summer research internship program with an emphasis on minority and women students was designed. Undergraduates were selected, as opposed to graduate students, for several reasons. First, we hope to recruit outstanding students in the physical sciences and mathematics and give them a positive research experience that will encourage them to attend graduate school and pursue research careers, Second, a survey of current NASA programs for students indicated that more opportunities were available for K-12 and graduate students than undergraduates.

The primary objective of the HSCaRS Summer Enrichment Program (SEP) is to make significant contributions to the NASA Mission to Planet Earth (MTPE) and the Alabama A&M University (AAMU) Center for Hydrology, Soil Climatology and Remote Sensing (HSCaRS) research missions by providing undergraduate student research internships with an emphasis on minority and women students. Additional objectives are to encourage more minority and women students to pursue advanced degrees in Earth system and global change science and to increase the participation of minority institutions in the U.S. Global Change Research Program. Also, the SEP strives to make students in the traditional science disciplines more aware of the opportunities in Earth System Science.

In designing the SEP, it was acknowledged that HSCaRS was a new research effort and Center. Consequently, students were not expected to immediately recognize the Center as one would older, more established research laboratories with national reputations, such as Los Alamos, Battelle, National Consortium for Atmospheric Research (NCAR), etc. Yet we still wanted to compete nationally for the best students. Therefore, we designed the program with a competitive financial package that includes a stipend of $400 per week, round-trip transportation from home to the summer research site, and free campus housing and meal plans provided by Alabama A&M University. Students also received a modest living allowance of approximately $25 per week. The internship program was 10 weeks in residence at Alabama A&M University or IGCRE, and gave students the opportunity to select from six general research areas: micro-meteorology, soil data analysis, soil moisture modeling, instrumentation, geographic information systems, and computer science. Student participants also enrolled in an introductory global change science course as part of the summer program (a copy of the course outline is in the appendix). The program included participation in a field program for approximately two weeks. All students were required to participate in the field program as a learning experience, regardless of the relationship of the field program to their majors or particular research project.

Recruiting and Evaluation Methodology

In the inaugural year of the SEP, students were recruited by distributing posters throughout the university community. All the HBCU's, Other Minority Universities (OMU's), and a mixture of other types of institutions

were targeted for the distribution of program information. Other institutions included a mixture of large schools with established research programs and smaller institutions with a traditional emphasis on teaching as opposed to research. Students that would complete the sophomore year prior to the summer internship period were targeted in the recruiting process. A total of 102 students applied for the program, and 10 were selected. The applications included 81 from minority institutions, of which 57 were minority women. By class standing, applicants included 30 sophomores, 52 juniors, and 20 seniors.

A two-phase process was used in the evaluation of the students' applications. The first phase was an evaluation by the SEP staff to identify the strongest applications by research group. This was done by considering the applicant's Grade Point Average (GPA), statement of interest in global change science, letters of recommendation, and resume. A grading scale was developed, that awarded two points for a GPA over 3.5, one point for a GPA between 3.0 and 3.5, and zero points for a GPA less than 3.0. An additional point or fraction of a point was awarded based on the staff's evaluation of the statement of interest in global change science, letters of recommendation, and resume. Consequently, these non-GPA evaluation factors provided the opportunity for student applicants to obtain three additional points. A maximum score of five points could be achieved.

This evaluation plan allowed us to consider all major relevant factors and was weighted toward students with the best grades. However, the plan did provide an opportunity for a student with a low GPA to still score high enough to be selected. This was considered important, since talented undergraduates sometimes do not achieve high grades early in college due to immaturity, changing interests, etc. Some of these students may still be very interested in research and have the talent to succeed in a research environment, The statement of interest in global change science offered insight into some outstanding candidates for our program that did not have high GPA's. Finally, relevant work experience as shown in the resume was considered to be of value.

Based on the phase one analysis, the students were ranked by score. Applicants scoring 3.5 or greater on the 5.0 scale, approximately forty percent of the total applicants, were considered further in the second evaluation phase. Phase two provided an opportunity for the potential mentoring scientists to rank in priority order the student applicants interested in the mentor's area of research. This data was used, with program goals in mind, to make final selections. The program goals that most prominently influenced the evaluation process related to the identification of students who would likely be able to make research contributions and who would be interested in graduate school.

Results

Overall, the first year was very successful. The program objectives for the first year were as follow:

1. Recruit five outstanding undergraduate students to pursue NASA and AAMU research projects.
2. Recruit, as a part of the five undergraduate students, a minimum of two minority woman students.
3. Have a minimum of two students make significant research contributions and return for a second summer.
4. Recruit, as a part of the five undergraduate students, a minimum of two students from predominately minority institutions.

In regard to the recruiting objectives, we were able to select 10 outstanding students, with the following demographic makeup: nine females and one male. Seven of the students were African-Americans, and three Caucasians. The average GPA of the students selected was 3.73 on a 4.0 scale, and the range was from 3.33 to 4.00. Student majors included Physics, Chemistry, Computer Science, Geography, and Environmental Science. This clearly meets the first two objectives.

At the end of the summer term, each student was required to prepare a written paper on his/her research and present their results in a seminar setting. Based on this data and feedback from mentoring scientists, all of the students made contributions to the research effort. Four of the ten students have been invited back for a second summer in 1997. Three of the four students are resident at HBCU's. We anticipate all of them will accept our offer and continue their research projects in 1997 that will lead to publications. Of these four students, one is continuing to work on the HSCaRS research remotely and may have a co-authored publication with an Alabama A&M University mentor prior to next summer, and one is working toward enrollment in graduate school at Alabama A&M University. All SEP students have been encouraged to attend professional conferences and present HSCaRS

research results. In particular, students have been encouraged to present their HSCaRS research at the NASA University Research Centers (URC) Technical Conference and the National Conference on Undergraduate Research (NCUR). HSCaRS has committed to paying all travel expenses for those students attending these conferences. A listing of student research project titles is included in the appendix.

The final year one objective was to recruit at least two students from minority institutions. This objective was achieved by recruiting seven students from Historically Black Colleges and Universities (HBCU's), as shown in the table below:

Home Institutions for 1996 SEP Students

Institution	Location	Classification	# Students
Alabama A&M Univ.	Normal, AL	HBCU	2
Carleton College	Northfield, MN	Majority	1
Fayetteville State Univ.	Fayetteville, NC	HBCU	1
Jackson State University	Jackson, MS	HBCU	1
Miles College	Birmingham, AL	HBCU	1
Norfolk State University	Norfolk, VA	HBCU	1
Spelman College	Atlanta, GA	HBCU	1
University of Maryland	College Park, MD	Majority	1
University of Oklahoma	Norman, OK	Majority	1

Source: SEP Program Office Records

Plans for 1997

Our plan for year two is to follow the successful script from the first year, with changes integrated from lessons learned. Due to the first year's success, the program is on track to meet or exceed all of the originally proposed year two SEP milestones. The original SEP milestones for year two are as follow:

1. Assuming two students return to continue research projects from the first summer, recruit three more outstanding undergraduate students to pursue NASA and AAMU research projects.
2. Recruit as a part of the three undergraduate students, a minimum of one minority woman student.
3. Have two undergraduate students co-author papers with mentor scientists
4. Have two undergraduate students enroll in graduate programs relating to Earth system and global change science.
5. Recruit as a part of the three undergraduate students, a minimum of one undergraduate student from a predominately minority institution.

Given the outstanding response by students to last year's program advertisement, it is reasonable to assume that the program is on track to meeting all recruiting related objectives. Our plans are to recruit ten new students, plus to invite back for a second summer, four outstanding students from the SEP class of 1996. Also, of the four students invited back for a second summer, at least two student co-authored papers are likely. The enrollment of one SEP student from the class of 1996 in graduate school to study in the computer science area at Alabama A&M is expected. Overall, the program is in a good position to meet or exceed all year two milestones.

Observations and Conclusions

First, it was exciting to see the overwhelming positive response to the 1996 SEP. The program was advertised primarily by posters that were distributed late in 1996 and early 1997 to several hundred institutions. Given the lateness and amount of the publicity, the receipt of over 100 applications and approximately twice that many inquiries about the program, is indicative of the interest in HSCaRS research and of the tremendous need for more undergraduate research opportunities.

The approach to the field program was that it would be beneficial to all the students to participate in this phase of the project to reinforce understanding that science is not done only in books. While the SEP program staff realized the field work would be more relevant to some student research projects than others, we did not communicate this effectively enough to the students. As a result, several students complained that there was not a direct connection between their research and the field program, and that this was time that could have been better spent on their respective research project. We will spend more time this year communicating program expectations and goals so similar frustrations will not occur.

About the midpoint of the summer term, we began to see a **correlation** between student performance in the introductory class in global change science and the students' background in Physics. Those students with at least one semester of Physics found the course to be much easier than the other students. As a result, we will emphasize the need for future student interns to take a course in Physics prior to beginning the summer internship. One anomaly to this conclusion was a Geography major who did very well in the course and with her research. Upon discussion of this issue, it was found that fundamental Physics Laws and basic theories had been integrated into other science courses as needed to fully explain and illustrate topics. Therefore, her Physics background was in fact sufficient. This is one illustration of why it is good to look beyond course titles and associated grade point average when evaluating students applicants. We think the evaluation methodology described above continues to give us the foundation needed to make the thorough analyses needed to ensure that we select the best possible applicants.

In regard to selection, global change science does encompass in some fashion a large number of courses of study. Consequently, one of our challenges was how much priority to give outstanding students in marginally related areas of study and interest in the evaluation process. We did have one example of a student who was mismatched in terms of project opportunities in HSCaRS and personal research interests. Nevertheless, this student was able to pursue a research project with some connection to HSCaRS research objectives, and overall had a productive summer. Conversely, the one student we recruited from the computer science area completed an excellent research project and as a result she is now planning to attend graduate school at AAMU.

Acknowledgments

Many thanks to the Center for Hydrology, Soil Climatology, and Remote Sensing (HSCaRS) mentor scientists, Phyllis Campbell, and Kate Hinke for assistance with this project. Contribution from HSCaRS, Department of Plant, Soil, and Animal Sciences, and the Agricultural Experiment Station, Alabama A&M University, Normal, AL 35762. Journal No. 344. This work was supported by Grant No. NCCW-0084 from the National Aeronautics and Space Administration (NASA), Washington, DC. Any use of trade, product or firm names is for descriptive purposes only and does not imply endorsement by the U.S. government.

References

The Task Force on Women, Minority, and the Handicapped in Science and Technology, "Changing America: The New Face of Science and Engineering", Interim Report, Washington D. C., September 1988, p. 3.

Feinstein, Selwyn, "Summer Internships Flourish, with Emphasis on Minorities", The Wall Street Journal, April 24, **1990**, front page,

Office of Space Science and Applications in Partnership with the Educational Affairs Division, "Looking to the Future: 1991 Catalog of Educational Programs and Activities", National Aeronautics and Space Administration, Washington, D. C., 1991.

Discussions with 1996 SEP Students and HSCaRS scientists.

Appendix

1996 SEP Student Research Project Titles

Determining Iron and Manganese Concentrations in Soil Using Radiometric Reflectance Readings

Student: Carrie Kienenberger Mentor: Dr. Ahmed Fahsi/AAMU

Physical and Chemical Characterization of the Research Test Bed at the Winfred Thomas Agricultural Research Station

Student: Tomeka Prioleau Mentor: Dr. Andrew Manu/AAMU

Assessing Rooting Traits in the Loblolly Pine and the Slash Pine to Compare Genetic Variations

Student: Latousha Parker Mentor: Dr. Ahmed Fahsi/AAMU

Using Ultrasonic Techniques to Measure Soil Moisture

Student: Kimberly Williams **Mentor: Dr.** Mohan Aggarwal/AAMU

Optical Sensing of Soil Moisture

Student: Mario Thomas Mentor: Dr. B. R. Reddy/AAMU

Creating a Cheap Model of the Relationship Between Soil and Hydrological Processes

Student: Barbara Cosgriff Mentor: Dr. Jason Kinser/AAMU

The Role of Topography in Water Movement and Energy Exchange at the Land Surface

Student: Ann Zawistoski Mentor: Dr. William Crosson/IGCRE

Simulation of Drainage, Water Content and Runoff of Four Varied Soil Types Using SHEELS

Student: Lucretia Jones Mentor: Dr. William Crosson/IGCRE

Patterns of the Radiation Balance as Influenced by Soil Moisture, Vegetation Cover and Meteorological Factors

Student: Malinda Taylor Mentor: Dr. **Charles Laymen/IGCRE**

Application of Modular Modeling System to Predict Evaporation, Infiltration, Air Temperature and Soil Moisture for Two Bare Plots

Student: Latrica Birgan Mentor: Dr. Teferi Tsegaye/AAMU

SPS 366- Climate and Global Change

4 Semester Credits
Lecture -3 hours weekly
Laboratory - SEP Research Project

Introduction to climate and global change including: the relationships between the Sun and the Earth that drive the climate system; the global structure and variations of the atmosphere and oceans, and the influence of humans and natural processes on the climate system and its variability. Additional topics include: the greenhouse effect; ozone depletion; air pollution; acid rain; biodiversity; paleoclimatology, and volcanism,

Topics

Introduction to Climate and Global Change
Atmospheric Variables and Measurements
Surface Variables and Measurements
The Earth's Radiation and Energy Budget
Remote Sensing Measurements
Atmospheric Motion and Global Circulations
Climate and Ocean Currents
The Earth's Hydrologic Cycle
Climate and Global Change
The Earth's Carbon Cycle
Interactions and System Dynamics
'Climate and Global Change

Source: Dr. Donald J. Perkey, Instructor

URC97042

/// ... 542 / 61

Integrity Constraint Monitoring in Software Development: Proposed Architectures

Francisco G. Fernandez*
Department of Computer Science
The University of Texas at El Paso 79968
ffernand@cs.utep.edu

1 Introduction

In the development of complex software systems, designers are required to obtain from many sources and manage vast amounts of knowledge of the system being built and communicate this information to personnel with a variety of backgrounds. Knowledge concerning the properties of the system, including the structure of, relationships between and limitations of the data objects in the system, becomes increasingly more vital as the complexity of the system and the number of knowledge sources increases. Ensuring that violations of these properties do not occur becomes steadily more challenging. One approach toward managing the enforcement of system properties, called *context monitoring* [2, 3], uses a centralized repository of integrity constraints and a constraint satisfiability mechanism for dynamic verification of property enforcement during program execution.

The focus of this paper is to describe possible software architectures that define a mechanism for dynamically checking the satisfiability of a set of constraints on a program. The next section describes the context monitoring approach in general. Section 3 gives an overview of the work currently being done toward the addition of an integrity constraint satisfiability mechanism to a high-level program language, SequenceL, and demonstrates how this model is being examined to develop a general software architecture. Section 4 describes possible architectures for a general constraint satisfiability mechanism, as well as an alternative approach that uses embedded database queries in lieu of an external monitor. The paper concludes with a brief summary outlining the current state of the research and future work.

2 Background

This section provides an overview of the context monitoring approach to software development. In addition, a detailed examination of a dynamic constraint satisfiability mechanism is provided.

2.1 Context Monitoring: A Brief Overview

Context monitoring [3] is an approach to software development that uses knowledge of the data objects of a software system to ensure program correctness with respect to selected properties. The approach consists of two parts [4]:

- the elicitation and specification of constraints on data or objects being modeled by the system, and
- a constraint satisfiability mechanism that dynamically verifies constraint enforcement during program execution.

Constraints on the data objects of a system can be identified at any stage of the development cycle. Domain experts and end-users identify constraints that define the behavior of the system. System developers may further refine the behavior of the system by making assumptions about the properties of data objects and the environment in which the program will run during the design and implementation of the system. For example, consider a software system for pharmaceutical sales. A government pharmaceutical board may

* THIS WORK WAS SPONSORED BY NASA UNDER CONTRACT NAG-101 2 AND NCCW-0089

state that each drug sold at the pharmacy is identified by a unique ID code specified by the board. In turn. the developers of the system may make an assumption about these ID codes, such as assuming that all ID codes are at most 15 characters in length. Such an assumption may not necessarily have been identified by the domain experts in field of pharmaceuticals, but is nonetheless a restriction on the operation of the system imposed by the developers.

Work is currently being done to define methodologies for eliciting system constraints from domain experts and cud-users during the requirements analysis, functional specification and other stages of software development (see, for example, [5]).

2.2 Constraint Satisfiability Mechanism

The constraints are specified as statements in first-order logic and maintained in a central repository. During program execution, this repository is consulted to deter-mine if the system is in fact enforcing the constraints.

An important, concept of the constraint satisfiability mechanism is the *state* of the program: a set of program variable-value pairs that capture a snapshot of memory at a given point of time during program execution. Constraint checks are based on changes in the state of a program.

The idea behind constraint satisfiability is to monitor programs for violations of constraints during execution. The state of the program is monitored and, when a change in state occurs, all constraints relevant to this state change are checked for violations, For example, if the value of variable X changes after an execution step, then only those constraints associated with variable X are checked for violations. If no violations occur, program execution continues. If a constraint violation does occur, the integrity of the program data has been compromised and as a result, program correctness from that point on can no longer be ass u red.

3 Constraint Satisfiability in SequenceL

This section describes the implementation of a constraint satisfiability mechanism in a high level language called SequenceL. A brief description of the language is provided, and the motivation for using it for the initial attempt is described. A description of the implementation is then given.

3.1 The SequenceL Language

SequenceL is a high-level language for processing non-scalar data [I]. In SequenceL, problems are solved by specifying the form and content of the data to processed. The iterative/rec ursive details of the problem are abstracted from the solution.

The basic data structure of the language is the sequence. From this simple structure, any other data structure can be constructed (i.e., sets, trees, records, etc..). In addition to the basic data structure, there are also basic operations which can be performed on sequences, such as addition, subtraction, multiplication, division, and other more complicated operations. A more detailed description of the language can be found in [1].

The language uses the notion of a *universe*, a collection of variable-value pairings. There are no explicit methods for input and output to a SequenceL program. Instead, program input is provided implicitly as the initial universe, and the final universe is the implicit output. A program begins with a collection of functions and an initial universe. A SequenceL function is defined by the domain arguments it must have in order to execute, the range arguments that arc produced as a result of processing the domain arguments, and the operations to be performed to process the domain and produce the range. SequenceL operates on an event-based model. As a result, function execution is not sequential, but based on the availability of parameter data in the universe. This makes the language non-deterministic, and also open to parallelizability. A. function is executed only if al] of the domain arguments of the functions arc available in the universe. [f a function is eligible, it consumes the domain arguments from the universe, executes the body of the function, and places the range arguments that result, from processing back in the universe.

3.2 Implementation of Constraint Satisfiability

Extending the SequenceL execution model for constraint satisfiability involves two major tasks. First, identifying where in the execution model constraint checks should be made. A state change occurs at a clearly identifiable point, i.e., at the end of execution of a function, The second more difficult task is to extend the

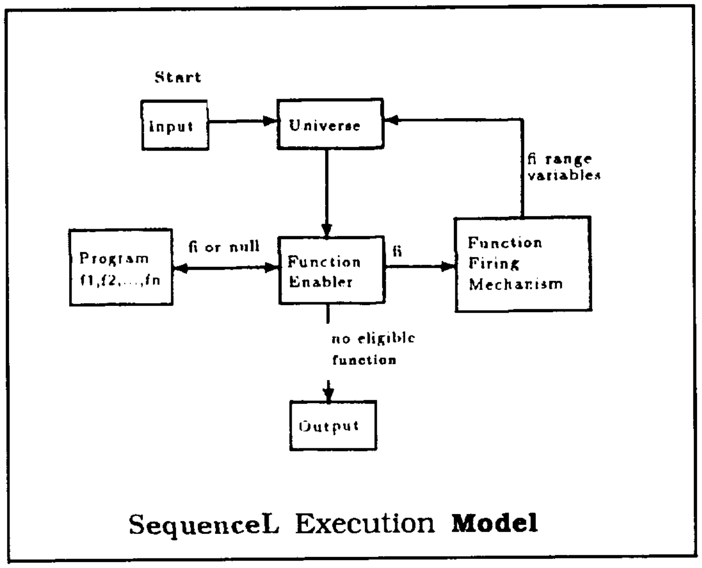


Figure 1: The SequenceL Execution Model.

SequenceL language to define what it means to check integrity constraint satisfiability for a program and a set of constrain. For simplification, the integrity constraint repository is represented by a program composed of boolean functions defining the relations and limitations of data objects in the SequenceL program. These constraint functions are defined as SequenceL functions, and thus the same execution engine is used to check the satisfiability of a constraint as is used to execute the SequenceL program. Figure 1 demonstrates the SequenceL execution model, and figure 2 shows the extended model with constraint satisfiability monitoring.

The development of a constraint satisfiability monitor in SequenceL is more than just an exercise. in fact, the reason for developing the monitor for a specific programming language is to study the interaction between constraint monitoring and program execution. The SequenceL model demonstrates how monitoring occurs in a program, and more importantly, provides a basic architecture model which can be used as the basis for a more general monitor.

4 Generalizing the Satisfiability Mechanism

The ultimate objective of the research is to create a general constraint satisfiability mechanism that can be used to ensure enforcement of data object properties independent of the programming language used to build a software system. This section examines possible software architectures for such a mechanism and alternatives to an external monitor.

4.1 Issues

To understand how to provide a general constraint satisfiability monitoring mechanism, it is important to first identify some of the issues.

Any general mechanism will require recta *knowledge* about the environment it will be monitoring, such as how variables take on values and knowledge that ties constraints to the program. As a simple example, consider a language that allows the use of array data structures. If the array is being processed within a loop, it might be the case that array should not be considered "changed" until the end of the loop processing. In other words, even though the array is undergoing a change during each iteration of the loop, it is not considered to be processed until the loop terminates. This type of information would be captured in the meta-knowledge.

A second consideration is how the monitor is related to the program it is monitoring. One possibility is that it forms a "wrapper" around the program, In this sense, the monitor is external to the program, and monitors it from the outside. It is somehow made aware that a state change has occurred in the program and that constraint checks must be made. Another possibility is that the monitor uses the meta knowledge

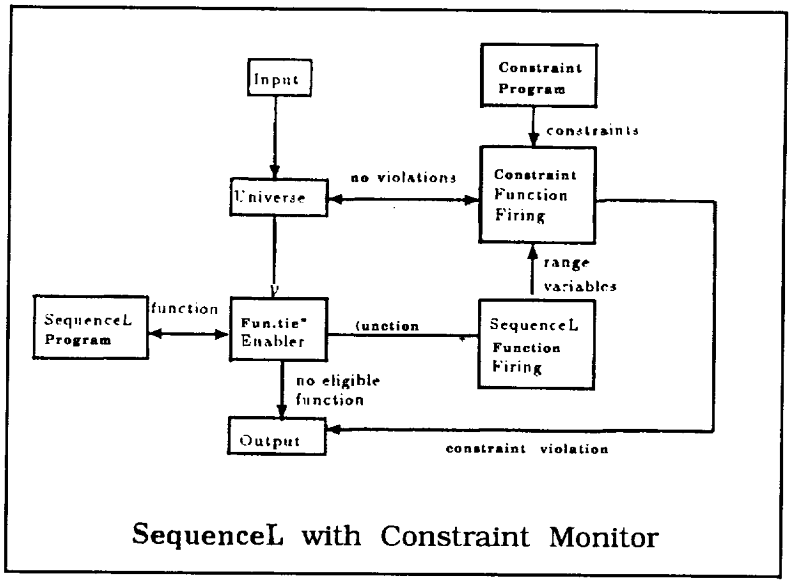


Figure 2: The SequenceL Execution Model with Integrity Constraints.

of the program environment to embed the constraint checks within the program itself. In this sense, the monitor is more of a pre-processor that transforms the program in such a way that it is able to monitor for constraint violations itself.

A third important issue is how constraints will be represented. This issue is dependent upon the implementation of the monitoring system. If the monitor is an external system independent of the program it is monitoring, it is possible to use a first-order logic representation of constraints. If constraints are to be embedded within the program it will be necessary to translate them from their first-order representation into equivalent representations in the language of the program

4.2 Architectures for External Monitoring of Constraint Satisfiability

Two architectures are outlined: the *event model* and the *tagged program model*. In the event model [7], changes in a program state implicitly trigger constraint satisfiability checks. The program broadcasts a change in state. The satisfiability mechanism, an external system monitoring the program, registers this state change broadcast and invokes appropriate constraint, checks. The monitor must use meta-knowledge to identify how the program will broadcast a change in state and how it will have access to the variables it must check for violations. Figure 3 gives a pictoral view of the event model, One advantage to the event, model is that the constraint monitoring can be easily parallelized. The program simply broadcasts that a change in state has occured and continues execution while the monitor checks the satisfiability of the constraints.

A second model is the tagged program model. Here, the monitor uses the recta-knowledge about the programming language to tag the source code for constraint checks. Prior to compilation, the program source code is passed to the satisfiability mechanism. The mechanism uses knowledge about the language and state changes to tag the program at points where checks should be made. During program execution, a constraint check is heralded when a tag is reached. Tags may be attached to variables in the program's symbol table, indicating constraints should be checked when the value of that variable changes. Alternatively, tags may be placed after program statements, indicating that constraints should be checked upon completion of that statement. Figure 4 shows an overview of the tagged program model.

4,3 Embedded Constraint Enforcement

An alternative to an external satisfiability monitor is the use of embedded constraint satisfiability checks. Some languages, such as C, provide constructs to embed assertions in programs to check program properties [6]. The main disadvantage to using embedded checks such as these is precisely the fact that the constraint checks are embedded. With embedded checks, it is not possible to reason about or study constraints outside

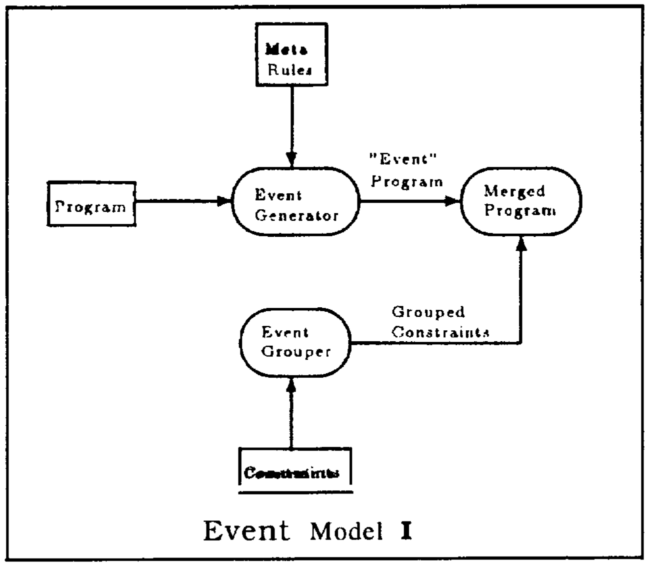


Figure 3: Architecture Diagram for the Event Model.

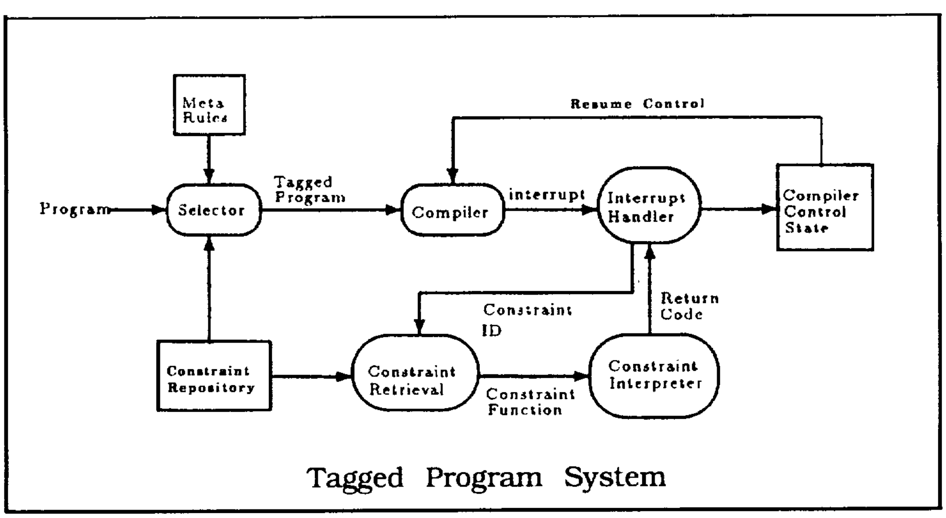


Figure 4: The Tagged Program Architecture Diagram.

of the program. If a change in constraints is required, it is necessary to untangle the embedded checks from within the program, which may be a difficult and tedious process.

Another possibility is to use language **extension** packages that provide interfaces to external querying systems. An example of such a package is EQUEL (Embedded QUEL), which provides FORTRAN with an interface Lo the INGRES relational database system. The disadvantage here is that every language would have to have an interface package specific to that language. Thus, the constraint satisfiability would not be a general mechanism, hut, a distinct one for each language.

5 Summary

Context monitoring and constraint satisfiability are powerful tools for managing the enforcement of vital system properties. By providing a system where constraints are maintained separately from a program it is possible to study, reason about, and modify them separately from the system. In addition, it may be possible to parallelize the constraint monitoring process to improve overall performance of the system.

Work is currently being be done toward designing a general satisfiability monitor. One issue being addressed is an analysis of programming languages to develop meta-rules about how variables under-go changes and how constraints can be associated to variables within a program. Also, methods for integrating a general constraint satisfiability monitor and different programming languages is being examined.

References

[1] Cooke, D. E., "An Introduction to SequenceL: A Language to Experiment with Constructs for Processing Nonscalars," to appear in *Software Practice and Experience,* 1996.

[2] Gates, A., *Context Monitoring With Integrity Constrain.* Las Cruces, NM: New Mexico State University, 1994 (Ph.D. Dissertation).

[3] Gates, A.Q. and Cooke, D. E., "The Use of Integrity Constraints in Software Engineering", *SEKE '95 Proceedings Software Engineering Knowledge Engineering,* Rockville, MI), 1995. Skokie, IL: Knowledge Systems Institute, 1995, pp. 383-390.

[4] Gates, A. Q., "On Defining a Class of integrity Constraints," *SEKE '96 Proceedings Software Engineering Knowledge Engineering,* Lake Tahoe, NV, 1996. Skokie, IL: Knowledge Systems Institute, 1996, pp. 338-344.

[5] Gates, A. **Q.**, "Building Systems with Integrity Constraints" ,*Proceedings of the Second World Conference on integrated Design and Process Technology,* Austin, 'TX, 1996.

[6] Rosenblum, D. S., "A Practical Apporach to Programming with Assertions," *IEEE Transactions Software Engineering, 21(1),* 1995, pp. 19-31.

[7] Shaw, M. and Garland, D., *Software Architecture: Perspectives* 071 an *Emerging Discipline* Upper Saddle River, New Jersey: Prentice Hall, 1996.

343/18

URC97043

NASA, WE HAVE A VISION: A Proposal to Launch a NANURC-Coordinated Tropical Orbiting UNEX Class Satellite Program to Mitigate Flood Disasters

Rafael Fernández-Sein
Tropical Center for Earth and Space Studies
University of Puerto Rico at Mayagüez

Abstract: NASA announced at the Vision'96 Conference a new class of satellites, denominated the University Explorer (UNEX), open to minority institutions. In this white paper we propose that the institutions represented in NANURC collaborate in the specification, design, construction, and operation of a satellite data collection and analysis network to produce early warnings of impeding flood disasters, This will be achieved by the fusion of data obtained by Synthetic Aperture Radars operating from space to determine accurate drainage basin parameters and flood stages with data obtained from the proposed satellite to estimate storm precipitation. This project will enhance the capabilities of each of the Alliance members, promote communications and linkages among the institutions, it will address some of the fundamental issues poised by the NASA mission statement, and will produce abundant international goodwill toward NASA and the US,

Introduction

NASA announced at the Vision'96 Conference a new class of satellites, denominated the University Explorer (UNEX), open to minority institutions. In this white paper we propose that the institutions represented in NANURC collaborate in the specification, design, construction, and operation of a satellite data collection and analysis network to produce early warnings of impeding flood disasters.

Inhabitants like me of the intertropical convergence zone know full well from experience that the Tropics are the heat engine that drives worldwide weather and climate. The Tropics, that region of planet 'Earth comprised within the Tropic of Capricorn and the Tropic of Cancer, or maybe between the parallels of 30°S and 30°N latitude, are the source of El Niño or Southern Oscillation (ENSO) and of the greatest heat engines: the enormously powerful storms known as hurricanes in the Atlantic and typhoons in the Pacific. How these gigantic forces of nature arise and interact are part of the questions that NASA seeks to find with its Mission To Planet Earth Enterprise. But the National Alliance of NASA University Research Centers (NANURC), can do much more than study these phenomena. This paper will out] ine a plan by which NANURC can collaborate to <u>mitigate</u> the disasters that are periodically visited on places far and near, but which arise in the Tropics.

Every year thousands of deaths worldwide are caused by flash floods. During its path through Puerto Rico during 10-11 September 1996, Hurricane Hortense brought such intense rains that 21 deaths were attributed to floods. Many of these deaths could have been prevented through adequate planning. The technology is now with us that can

provide accurate forecasts in sufficient time for evacuation of populations that are in danger of storm surges, flash floods, or long-term seasonal floods, thus mitigating the effects of such disasters. Let us review the methodology by which these can be effected.

Concept Exploration

At present, there are a multitude of satellites aloft, many of them devoted to meteorological concerns. There are the GOES and Meteosat geostationary satellites, and the NOAA series in polar orbits. However, the geostationary satellites cannot give the resolution that the application at hand requires, and the polar orbiting satellites spend the majority of their orbits away from the Tropics. We therefore, need a satellite to orbit at a near-equatorial, low altitude, low angle of inclination orbit, This satellite we shall call the Tropical Explorer, or TropEx, for short. For the development of its mission we follow the outline shown in Larson and Wertz [1].

The TropEx payload of scientific experiments, sensors, and other characteristics should be decided by the NANURC institutions. Certainly, to be able to carry out the flood mitigation mission outlined above, the TropEx should carry sensors in the visible spectrum to track hurricanes and assess ocean and land colors, as well as infrared sensors to estimate temperatures at different levels of the atmosphere. However, the most important sensors that should be included are those for the estimation of moisture and possible precipitation at given area. The process by which the experiments will be chosen are outlined in the section below dedicated to mission scheduling. The concept for the TropEx Mission can be summarized below.

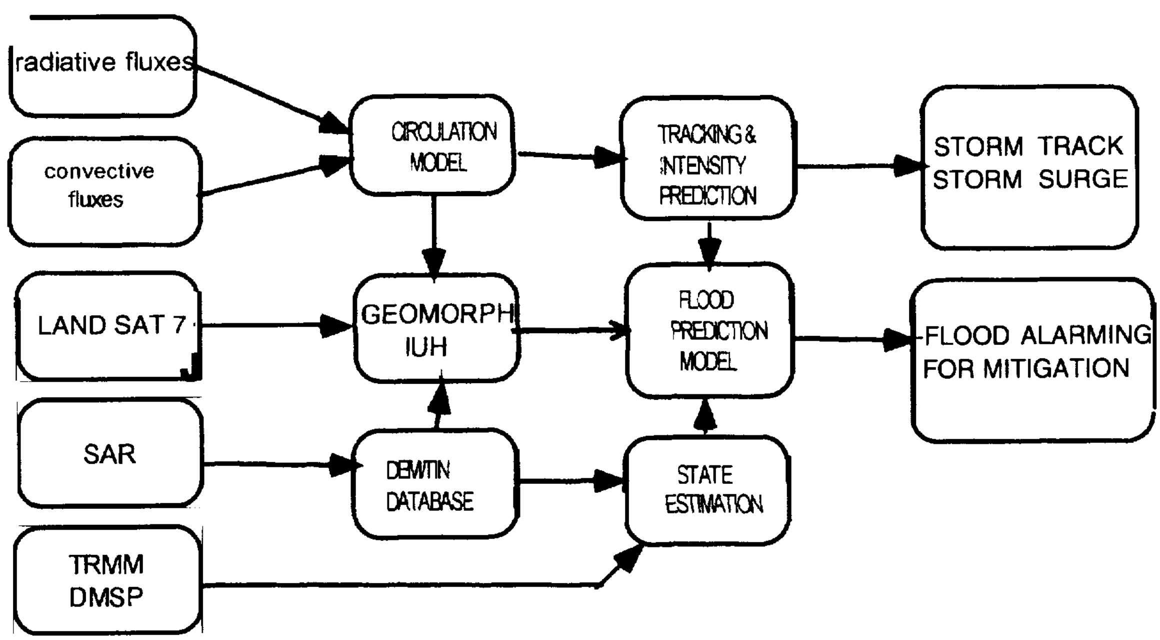


FIGURE 1. Conceptual flow graph of information for TropEX Mission.
Mission Objectives

To comply with the NASA mission statement, the NANURC UNEX should have clearly defined scientific objectives. These objectives should include measuring the parameters necessary to establish circulation models that will help in the understanding of hurricane formation and their forecasting and tracking. For the forecasting of floods, hydrologic models of the basins will have to be constructed. Finally, some means to measure atmospheric humidity and possible precipitation from space will have to be provided. All these measurements will have to be coordinated, analyzed, and reduced to forecasts that will allow mitigation of the disasters. But in addition to the scientific objectives, academic and social/industrial objectives. All these are summarized in the table below.

Table 1. Goals and Outcomes for the NANURC UNEX

SCIENTIFIC	ACADEMIC	SOCiAL/industrial
1. Understand the radiative and convective phenomena that lead to vortex formation, 2. Develop algorithms for the development of hydrologic models of the regions of interest. 3. Develop tracking algorithms for hurricane tracking and forecasting. 4. Develop algorithms for flood forecasting and alarming.	1. Provide training and education for our students at all levels: BS, MS, PhD. 2. Train students in every aspect of a satellite project, including design, modeling, fabrication, and testing. 3. Provide a focus for NANURC institutions and a means to collaborate. 4. Allow NANURC institutions to join mainstream institutions. 5. Encourage student research.	1. Furnish the Human Potential required by industry. 2. Develop products that can be marketed to industry. 3. Allow students to develop satellite-related enterprises.

Mission Utility

The TropEx Flood Mitigation Mission will bring together to bear on the problem experts from several disciplines, including meteorology, oceanography, fluid mechanics, atmospherics, hydrology, geomorphology, mathematics, mechanical engineering, electrical engineering, and computer engineering. All these experts will come from the NANURC institutions and from mainstream institutions willing to collaborate in this undertaking. A mapping of the collaboration areas, by no means final, or all-inclusive or definitive is presented below. Of course, this is presented in the spirit of finding a common ground for collaboration, and is shown as an example.

Table 2. Collaborations Matrix

	REM CTL	SPACE CRAFT	REM. SENS.	GND. STA.	ELEC TRO- NICS	HY- DRO- LOGY	FLUID MECH	SEN - SORS	HU- MAN FACT.
CLARK		x							
NC A&T		x			x				
TUSKEGEE			x			x			x
TENN. ST.	x								
AL. A&M			x			x			
FLORIDA							x		
HOWARD		X							
HAMPTON								x	
FISKE								x	
MOREHOUSE									x
PRAIRIE VIEW					x				
UNM		x			x				
UTEP			x	x					
PUERTO RICO	x		x	x	x	X	X	x	

Mission Planning, Scheduling, and Management

The TropEx mission must be planned a launch date within the time span of our grants. This means that the launch should occur on or before the year 2000. In a certain way this would be a most auspicious happening for our institutions, for it would mark our coming into the very mainstream of research and development. Below is presented a tentative timeline for this event to be achieved.

Table 3. Tentative Schedule

TASK NAME	Planned Start	Duration in Days	Planned Finish	1997	1998	1999	2000	2001	2002
PLANNING	1/1/97	270	9/28/97						
PROJECT DEFIN	9/28/97	270	6/25/98						
DESIGN	6/25/98	360	6/20/99						
CONSTRUCTION	6/20/99	360	6/14/0						
DEPLOY	6/14/0	90	9/12/0						
OPERATION	9/12/0	720	9/2/2						

Scientific Basis for the TropEx Flood Mitigation System

The theory for basin dynamics is summarized in Fernández-Sein [2]. We assume that effective rainfall is converted to direct storm runoff in a linear-time-invariant process. The unit hydrography $h_T(t)$, akin to the unit impulse response of circuit and systems engineers, has been defined as the unit volume of rain falling uniformly in a unit period T. The output flow time series for a catchment can then be written as the discrete convolution of the input time series i(kT) and the unit hydrography:

$$Q(nT) = \sum_{m=0}^{\infty} h_T(nT - mT)i(nT) \qquad n = 1, 2, \cdots \tag{1}$$

Nash [3,4] developed a conceptual instantaneous unit hydrography consisting of a series of α linear reservoirs, each emptying into the next reservoir in the series. The model is obtained by convolving α times the upstream inflow with the linear reservoir IUH of the type $h(t) = 1/\tau \exp(-t/\tau)$, where τ is the time constant for the reservoir. This leads to the two-parameter Nash model:

$$h(t) = \frac{1}{\tau}\left(\frac{t}{\tau}\right)^{\alpha-1} \frac{1}{\Gamma(a)} e^{-\frac{t}{\tau}} \tag{2}$$

The geomorphologic instantaneous unit hydrography (GIUH) provides for the estimation of the unit impulse response of a basin from quantitative morphology. The basic assumption for this methodology is that the GIUH can be expressed as a function of geomorphologic parameters, such as Horton's [5] order ratios R_A, R_B, and R_L, corresponding to the stream area ratio, bifurcation ratio, and stream length ratio,

In 1984 Rosso [6] modified the conceptual Nash formulation, turning it into a GIUH by defining the Nash parameters in terms of Horton's order ratios, a v mean streamflow velocity, and L a mean length of the higher order streams already mentioned. The Rosso model allows for the calculation of basin parameters from quantities easily determined from a GIS for the basin or from DEM data obtained from SAR imagery. The difficulty with this model is the dependence on the mean stream velocity v. Some authors have attempted to remove this restriction, but in doing so have complicated the model by making it dependent on the mean effective rainfall intensity, thus turning the model into a geomorphoclimatic instantaneous unit hydrography, Many other geomorphologic models have been developed. We have presented probably the simplest and most direct of the models available.

The computation of the various parameters required for geomorphologic analysis can be readily obtained by the use of appropriate data structures. Since the morphology of a river basin is basically a tree, with the "root" corresponding to the outlet and the "leaves" to the upper reaches. To place this information into a "smart" database, we store point information (location and elevation of river nodes, cross sections at potential flooding sites), arc information (length of river reaches), and area information. Areas are entered as polygon data, indicating the perimeter points, and interior points bounding possibly different land-use or land-cover polygons. The storage requirements for this type of data is drastically reduced by the use of triangulated irregular networks (TIN).

The required rainfall rate inputs that the GIUH models require can be obtained by multifrequency radiometric observation of the cloudstructures from a spaceborne platform. Pierdicca, Marzanno, et al. [7], for example, generate a series of databases of cloud structures and related brightness temperatures obtained from the Special Sensor Microwave/Imager (S SM/I) aboard the DMSP satellites. From these databases a Cloud

Radiation Database is combined with the Radiometric Multispectral Image an inverse problem to obtain the rainfall rate estimates.

As can be seen all of these processes and algorithms require gigantic databases, and a great amount of collaboration to digest and integrate the data streaming from the sensor platforms. However, the possibility is certainly there.

The TropEx Mission as a Focus for Collaboration

A mission with lofty objectives is a sure method to focus and generate the necessity of collaboration. However, what has been presented here is only an outline and suggestion as to what maybe proposed, with a consensus opinion to be determined by the NANURC partners. Means will have to used to enhance remote collaboration of groups, such as teleconferencing, and intensive use of the INTERNET. However, since each institution will be working within the limits of its expertise, the project is certainly possible and viable. What is now necessary is the will.

NASA, we have a Mission . . . .

Bibliography

1. Wiley J. Larson and James R. Wertz, editors, *Space Mission Analysis and Design,* Second Edition, Microcosm Inc. and Kluwer Academic Publishers, 1992, ISBN I-881883-01-9.
2. Rafael Fernández-Sein, "Design of a Flash-Flood Forecasting System", Proceedings of GIS/LIS'95, Nashville, Term., Nov. 14-16,1995
3. J.E. Nash, "The Form of the Instantaneous Unit Hydrography", International Association for Scientific Hydrology, Assembleé Générale de Toronto TOME 111, pp. 114-121, 1957.
4. J.E. Nash, "Systematic Determination of Unit Hydrography Parameters", J. Geophys. Res. 64:111-115, 1959.
5. R.E. Horton, "Drainage Basin Characteristics", Trans. Am. Geophys. Union, 13:350-361.
6. Rosso, "Nash model relation to Horton order ratios", Water Resources Research, Vol. 20, pp. 914-920, 1984.
7. N. Pierdicca, F. Silvio Marzano, G. d' Auria, P. Basili, P. Ciotti, and A. Mugnai, "Precipitation Retrieval from Spaceborne Microwave Radiometers Based on Maximum a posteriori Probability Estimation", IEEE Trans. On Geoscience and Remote Sensing, Vol. 34, No. 4, pp. 831-846, July 1996.

URC97044

5 44/20

On the Development of a Magnetically Vectored Variable I_{SP} Plasma Rocket

Enectalí Figueroa Feliciano
Stanford University; Stanford, California

Franklin R. Chang Díaz
Advanced Space Propulsion Laboratory; NASA JSC, Houston, Texas

Jared P. Squire
Advanced Space Propulsion Laboratory; NASA JSC, Houston, Texas

Abstract

The development of a Magnetically Vectored Variable I_{SP} Plasma Rocket at the Advanced Space Propulsion Laboratory (ASPL) is in progress at NASA's Johnson Space Center. The facility is using a small, 3.2 m tandem mirror device to study the application of RF heated magnetically contained plasmas for space propulsion. The central cell radius is 0.1 m and fields of 0.2 T and 2 T are possible in the central and end-cell mirror sections, respectively. A magnetoplasmadynamic (M PD) injector has just been acquired and will be used along with other methods of plasma refueling. A 1 M W magnet power supply upgrade is being developed with full implementation by the spring of 1997. Two microwave systems for discharge initiation and plasma heating at 2.45 GHz and 14.0 GHz, respectively, are in operation. Additionally, RF systems with 200 kW and 1 MW of power are being modified and conditioned for operation. The concept provides electrode-less operation and variable thrush'specific impulse at constant power (200 -30 N / 5000-30,000 seconds at 10 M W with a 60% efficiency). Optimization for speed or payload are possible with the same engine, giving the rocket great flexibility. Missions to Mars in 90 days are described, and missions to Pluto are under study,

Introduction

Human and robotic interplanetary travel using conventional chemical rockets suffer from prolonged trip time. In the case of human missions, the exposure to weightlessness and radiation and their effect on the crew's health are the major concern. The possibilities for changing trajectories in mid-flight are very constrained, and the total maneuverability once the destination is obtained is limited to correction burns and orbit captures. In the event of an emergency, abort scenarios in human missions are very difficult and often involve long orbits and close solar flybys. For robotic missions, the years of coasting through interplanetary space while principal investigators wait to begin

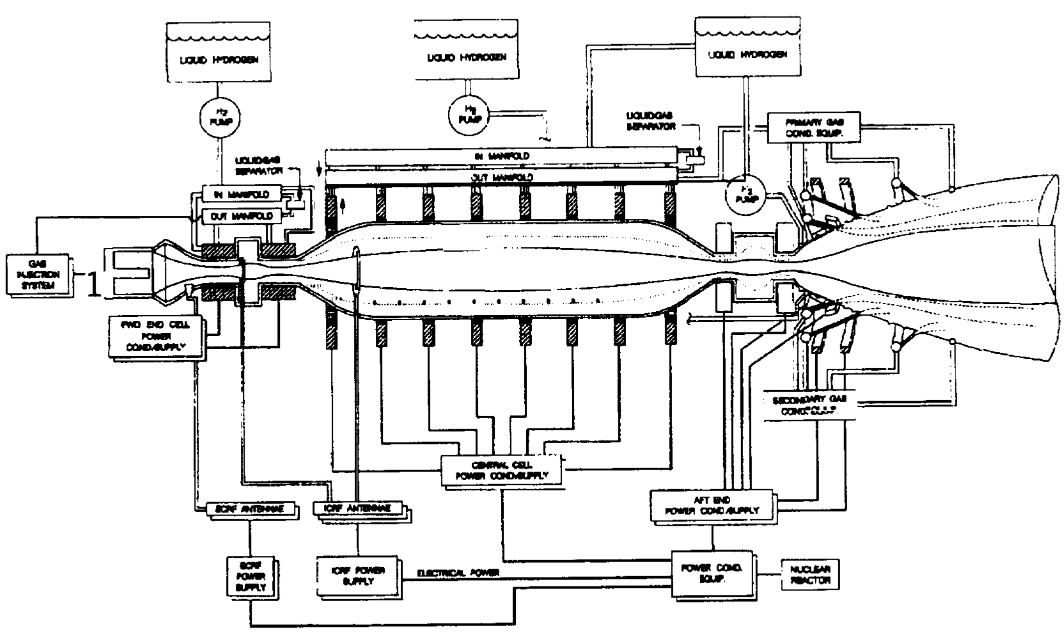


Figure 1. Schematic of the Magnetically Vectored Variable I_{SP} Plasma Rocket.

gathering their data makes involvement by scientists and students taxing, and raises the ground support and total cost of such missions. The acceleration time in conventional chemical rocket is negligible when compared to total trip time. This in turn puts constraints on the available launch windows, especially for long missions to the outer planets, reducing the versatility and error correction capability of the spacecraft. What is needed is an advanced propulsion system capable of optimization for high payloads, high speed, or maneuverability for the duration of the mission.

One such capable system is the Magnetically Vectored Variable I_{SP} Plasma Rocket.[1] This approach uses a tandem mirror configuration used in fusion research to magnetically contain a plasma heated by Ion Cyclotron Resonance Heating (ICRH). The high powered, electrode-less thruster can operate optimally over a large range of thrust/specific impulse (I_{SP}) combinations. This is accomplished via a variable magnetic nozzle with a coaxial gas injector near the throat that modulates the exhaust and protects engine materials from the hot plasma. A schematic is shown in Fig. 1.

Concept

A tandem magnetic mirror[2] consists of a collinear set of annular magnets divided into a solenoidal central cell, where plasma is contained and heated, and two higher field end mirror sections. These end-cell sections are high field magnets that tend to reflect the plasma at either end of the central cell and axially trap a fraction of it; thus the term "magnetic mirror." When used for propulsion, gas (typically an inert gas or hydrogen) is injected through one end and ionized into a cold, dense plasma by a device such as a MPD injector, a, hollow cathode discharge or by Electron Cyclotron Resonance Heating (ECRH) before entering the central cell, The central cell acts as an amplifier and heats the plasma by ICRH.[3,4] The device is operated asymmetrically, inducing a flow throttled by the exhaust nozzle magnetic field. At the nozzle, a thin, dense film of hypersonic gas is injected to aid the detachment of the charged particles from the magnetic field by diffusion, adding the benefit of insulating the engine materials from the hot plasma. Additionally, an AC component to the DC field maybe introduced at the nozzle to create instabilities and "shake" the plasma off the field lines. The magnitude of the AC component and the mass flow of the insulating discharge gas depend on the particular mode of operation, and are expected to be minimized at the low thrust, high I_{SP} mode.

Basic Equations

The advantages of variable I_{SP} propulsion have been known since the early 1950s.[5,6,7] The conceptual basics are developed here in the simplest but illustrative case of gravity-free space. For a variable I_{SP} rocket, the exhaust velocity with respect to a frame at rest with the vehicle departure point is made to match the vehicle velocity over the majority of the trajectory. Exhaust kinetic energy is optimally imparted to the vehicle when the exhaust leaves the rocket at rest, i.e., $u = v$, where the exhaust and vehicle velocities are denoted u and v, respectively.

Assuming no gravitational effects, the momentum balance for a rocket in free space is

$$\dot{v} = -\frac{u\dot{m}}{m} \tag{1}$$

where $\dot{v}$ and $\dot{m}$ are the time derivatives of the vehicle velocity and rocket total mass respectively, and m *is the* total mass. The total power flow off the exhaust is

$$P = \frac{u^2}{2}\dot{m} \tag{2}$$

and the relationships between the various masses are

$$m = m_o - m_f \tag{3}$$

where m_o *is the* total mass of rocket fuel at t = O, and m_f is the total mass of fuel exhausted through the nozzle at time t.

For a constant I_{SP} rocket, $u = I_{SP} * g$ = constant, and Eq. (1) can be integrated between **initial** and final velocities v and v_o to give

$$v(t) = v_o + u \ln \frac{1}{1-y} \tag{4}$$

where

$$\gamma(t) \equiv \frac{m_f}{m_o} = \frac{\dot{m}}{m_o} t \tag{5}$$

is the mass fraction as a function of time. The mass flow rate $\dot{m}$ is usually considered a constant.

For a variable I_{SP} rocket, if we optimize efficiency so that $u = v$, integration of Eq. (1) yields

$$v = v_o \frac{m_o}{m} = v_o \left(\frac{1}{1-\gamma} \right) \tag{6}$$

Equations (4) and (6) are plotted in Fig. 2. An exit velocity of $u = .5\, v_o$ was used for the conventional rocket, For the same mass fraction an optimally tuned variable I_{SP} rocket obtains a significantly greater velocity than a conventional rocket; for a given velocity, variable I_{SP} allows a far smaller mass fraction, thus granting larger payload capacity.

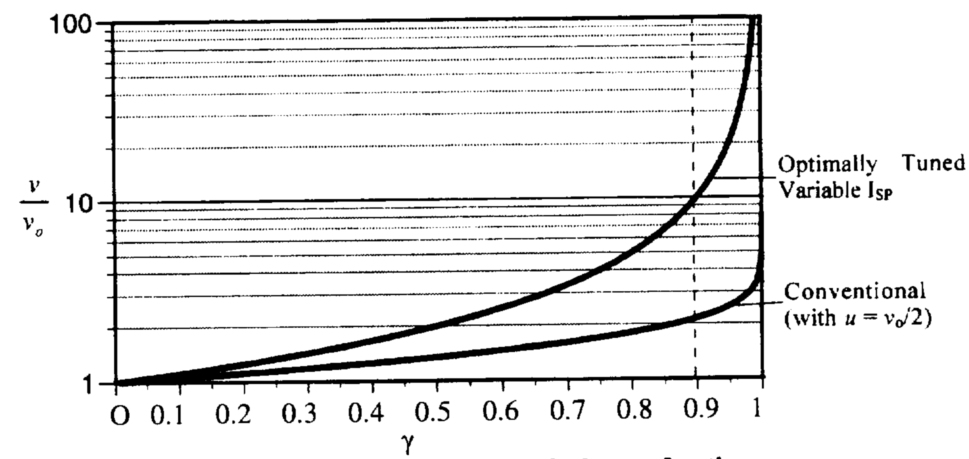


Figure 2. Velocity ratio as a function of the fuel mass fraction γ.

It should be noted that similar velocities can be achieved by high constant I_{SP} engines. However, when the distance involved is large,as in the case of interplanetary travel, variable I_{SP} reaches high speeds faster, and thus reduces the total trip time.[8]

Facility and Operation

To establish experimentally the feasibility and characteristics of the Variable I_{SP} Plasma Rocket, a tandem mirror plasma facility has been established at the Advanced Space Propulsion Laboratory (ASPL). The laboratory is located at NASA's Sony Carter Training Facility (SCTF), part of the Johnson Space Center in Houston, Texas. Plasma discharges have been created successfully at low energy levels.

The laboratory has dimensions of 96' x 48', The main experimental bed consists of a small tandem mirror machine, 3.2 m in length by 0.75 m in diameter. The vacuum chamber is a cylindrical assembly consisting of 9 stainless steel sections, with 24 ports available for diagnostics. A turbomolecular pump at the exhaust end easily maintains a 10-7 torr base vacuum. Back-filling to approximately 10^{-4} torr for plasma discharges is accomplished with a "10-100 seem gas flow. Our present pumping rate capability via turbomolecular pumps is 1,500 liter/s, although a cryopump is being considered that will boost the pumping rate to over 5,000 liter/s. An exhaust tank 1.5 m in diameter by 3 m in length has been fitted to the machine through a cone mating section and will permit high density discharges and study of the plasma plume in a simulated space vacuum. The cone section will later be an area of intense study as hybrid gas-magnetic nozzle experiments begin in 1997. The copper winding magnet set consists of two tandem mirror assemblies, each capable of 6,000 A each at liquid nitrogen temperatures, and the central cell, which has eight magnets capable of 1,000 A each at ambient temperature. To cool the end-cell magnets to liquid nitrogen temperatures, vacuum dewars have been designed into the machine; compressed air circulation

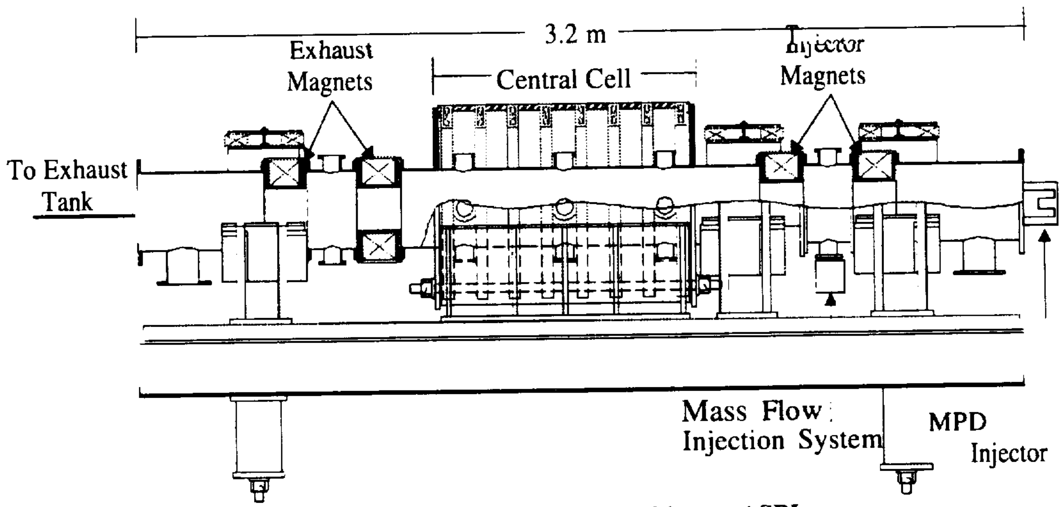


Figure 3. Tandem mirror machine at ASPL.

cools the central cell magnets. The magnets are presently powered by a 45 kW system consisting of three independent DC power supplies to regulate the current to each part of the machine and tailor the magnetic field for optimum flow, A 1 MW pulsed (10 seconds) DC power system is in the installation phase and should be operational by the time of this conference. Two microwave sources are used for ECRH: a 2.45 GHz microwave generator capable of delivering 800 W and a 2 kW, 14.5 GHz Klystron microwave generator that has been brought into operation this summer. For ICRH, two systems are currently being modified and conditioned for installation in the lab. One is a 200 kW RF transmitter, and the other is a set of three 1 MW RF transmitters that at one time powered the Arecibo Radio Telescope. They both have a frequency range from 1 to 10 MHz. For data acquisition, 5 CAMAC crates with a total of over 100 channels are connected to a data highway providing 2 Mbytes/s down loading rate into a Power Macintosh. A summary of the laboratory setup is shown in Table 1.

System Parameters

Overall length	3.2 m
Central cell length	1.15 m
Central cell radius	0.36 m
Exhaust tank	d-1.5 m x 3 m
Vacuum level	10^{-7} torr
Central cell field (B_c)	2 kG
Maximum field (B_{max})	3 T
DC power	1 MW
Plasma density (variable)	10^{19} m^{-3}
Plasma temperature	0,01 -1.0 keV
Plasma heating power:	
ECRH	2 kW
ICRH	1 M W

Table 1. System Parameters for ASPL equipment.

A brief overview of current operation follows. After the desired base vacuum is achieved (104 tom), a mass flow control system is used to inject Argon or Helium into the chamber in the injector end of the machine. By adjusting the flow, a pressure of 2x 10-4 torr is established, which is the optimum pressure for plasma breakdown in our current system. Three separate power supplies are used to selectively modulate the magnetic field in the injector, the central cell and the exhaust. Plasma discharges can now be generated in two modes, low and high power, For the low power mode, the 2.45 GHz generator is used to induce ECRH in the gas and generate a plasma. This mode, which produces a density of $7x1O^{16}$ particles/m^3, is used to test the equipment and purge the vacuum tank of impurities, since it can be run at steady state. Currents of 150 A in the injector and 30 A in the central cell magnets are typical for this mode, generating magnetic fields up to 0.1 T. The high power mode uses the 14.5 GHz Klystron in conjunction with the 2.45 GHz generator to achieve densities of 2x1O particles/m^3. This mode can be

operated only for 30 second pulses, and liquid nitrogen refrigeration is needed for the end-cells. Part of the activities of this summer included the design of a Programmable Logic Controller (PLC) system for automating the operation of the machine, which is now done manually.

Development

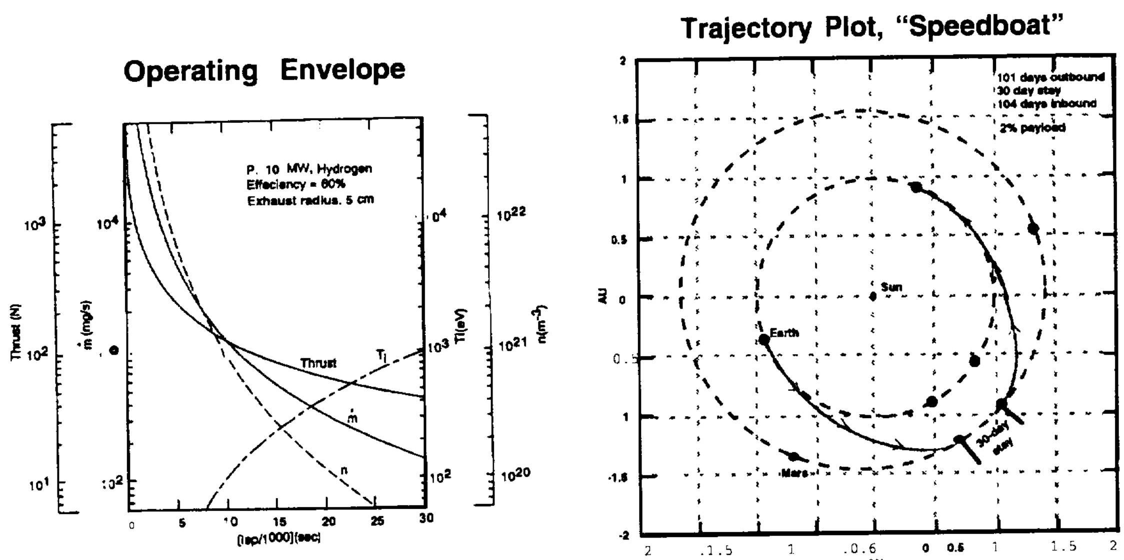
Several exciting prospects await ASPL in 1997. The completion of the system installation and fill on-line status in expected, at which point the major focus of the lab will shift into experiments and data collection. There are still numerous questions to be answered and concepts to be proven, and there are plenty of opportunities in engineering and plasma physics to follow. Several diagnostics will be installed, including an array of Langmuir probes to measure the temperature and density of the plasma at the edge. A laser fluorescence system at 6563 Å will determine the neutral density of the exhaust, and a microwave interferometer operating at 100 mW will be used to map the plasma density. Also, a retarding potential analyzer for the plasma exhaust is under development. Further studies will be conducted, especially in the nozzle section, where magnetic field decoupling effects have to be ascertained. A general empirical performance envelope for the concept must be developed, and scaling laws for different power regimes determined.

Proceeding into the next step of development, a small, 30 kW engine is being designed to fly as a shuttle mission in the coming years. Powered by batteries, it would be a free flying satellite deployed and controlled from the Shuttle and would perform many 1 to 2 minute pulses. It would carry superconducting magnets and its own communication system. The goal of the mission would be to prove the general principle of the concept and test the process in the vacuum of space. Thrust and I_{sp} would be monitored, along with a host of plasma diagnostics and video feed to have visual contact with the engine and plasma while in operation.

Next a 100 kW demonstration rocket is envisioned. This would be very similar to a full scale version, possibly nuclear powered, with superconducting coils. Once operational, it could be used to launch magnetosphere mapping missions in a outward spiral around Earth; or to take cargo, such a GPS transceivers or supplies, to the moon; or to send small probes to the outer planets in record times.

Future

The full scale Magnetically Vectored Variable I_{SP} Plasma Rocket would be a high redundancy, high power platform able to carry large payloads in a "tugboat" mode, or to safely and quickly carry a crew in a low payload "speedboat" mode. A recent study[9] examined scenarios for human and robotic missions to Mars. The mission is a split sprint consisting of a robotic mission that would carry supplies and landing modules, and a fast light mission carrying the crew. The same type rocket would be used for both, and would be a 10 M W version of the demonstration rocket



(a) (b)

Figure 4. a) Operating Envelope for a 10 MW rocket. b) 101 day human mission to Mars for a 30 day stay. Travel times of 90 days with 14.49% inbound payload capacity and 18.12°/0 outbound payload capacity are possible for extended duration (two year) missions,

mentioned in the last section (Fig. 4a). For the robotic mission, a 66.66% payload mass fiction and 180 day trip is projected, at which time the craft will enter Mars orbit. Check outs will be done and a robotic lander deployed with a habitat, fuel, and supplies in the two year wait for Mars to again be in optimum position for flight. Once the two year wait is over, the crew will leave Earth orbit for a 101 day trip. Upon entering Mars orbit, the crew will dock with the supply module, and descend to the Martian surface. Thirty days later the crew will be back in their speedboat en route to Earth (Fig. 4b).

Abort capability plays an important role in the feasibility of a mission scenario. During the Apollo program and currently in the Shuttle program, abort possibilities were and are inherent in mission design. Interplanetary missions should live up to and out-do the capabilities of these programs. Were the ship's life support systems to malfunction, or a crew member to become sick, or one or more of the modules of the rocket cluster to fail, the ship must have the capability to return to Earth, or at least arrive at the supply module orbiting Mars. For 90 day trips, if an abort condition occurred 15 to 20 days after leaving Earth orbit, a 75 day abort to Earth is possible with a full-up propulsion system. In the case of partial propellant loss, a 180 day abort is possible, although the amount of propellant and velocity away from Earth at abort time will determine the best abort trajectory.

After the first mission, the habitat that would remain in the Martian surface and the power plants left in orbit could be used by future missions. If a longer stay is desired, the crew would have to "winter down" on the Martian surface for two years. If a two year mission is planned, 90 day trips with 14.49°/0 inbound payload mass fraction are possible. Moreover, while still making the Mars-Earth trip in 90 days, the crew would have an 18.12% payload mass fraction upon return, allowing for a sample return equal to 3.63% of the mass of the ship.

Conclusion

The Magnetically Vectored Variable I_{SP} Plasma Rocket provides optimum thrust/I_{SP} modulation for fast interplanetary missions. Much theoretical and experimental work remains to be done to understand the optimum operating parameters and the behavior of plasma in an asymmetric tandem mirror. Although initial numerical simulations have been made,[10] the nozzle section and the processes which separate the plasma from the diverging magnetic field must be further studied. ASPL is devoted to developing this technology and experimentally demonstrating its feasibility in space borne missions in the coming years. The laboratory is undergoing major upgrades in the coming months and should be breaking new ground in power and data acquisition in 1997. Ongoing projects include a study into the trajectory for a Pluto flyby mission, continuing out into the heliosphere to study deep space to a the thousand astronomical unit (TAU) mark.

Mars, although the probable first planet in our human exploration of the solar system, will not be the last. Technologies must evolve and our commitment toward exploring the Solar system must solidify. Humans will need to master the art of living and working in space. To get there, we must continue to develop better and innovative propulsion systems. This is the goal of ASPL.

References

[1] Chang Díaz, F. R., "The Hybrid Plume Plasma Rocket," NASA Johnson Space Center Internal Report, January 1982.

[2] Balwin, D.E. and Logan, B.G., "TMX Major Project Proposal," Lawrence Livermore National Laboratory Report 11-Prop-148, 1977.

[3] Yang, T. F., Peng, S. and Chang Díaz, F.R. *AIAA/ASME/SAE/ASEE Int. Joint Propulsion Conference, paper AIAA-91-2338*, Sacramento, California, 1991.

[4] Peng, S., "ICRF Wave Propagation and Plasma Coupling Efficiency in a Linear Magnetic Mirror," Ph.D. Thesis, MIT, 1991.

[5] Irving, J.H. and Blum, E.K., "Comparative Performance of Ballistic and Low-Thrust Vehicles for Flights to Mars," *Vistas in Astronautics,* Vol II, Pergamon Press, New York, 1959.

[6] Stuhlinger, E., *Ion Propulsion for space Flight,* McGraw-Hill, 1964.

[7] Melbourne, W. G., "Interplanetary Missions with Low Thrust Propulsion of Advanced Propulsion Vehicles," Jet Propulsion Laboratory Internal Report 32-68, Pasadena, California, March, 1961.

[8] Chang Díaz, F. R., Yang, T.F., Kruger, W.A., Peng, S., Urban, J., Yao, X. and Griffin, D. *DGRL/AIAA/JSASS Int. Electric Propulsion Conference, paper DGLRA-88-126,* Garmisch-Partenkirchen, W. Germany, 1988.

[9] Chang Díaz, F.R., Hsu, M.M., Braden, E., Johnson, I. and Yang, T.F., "Rapid Mars Transits with Exhaust-Modulated Plasma Propulsion," NASA Technical Paper 3539, March 1995

[10] Krueger, W.A, "Plasma and Neutral Jet Interactions in the Exhaust of a Magnetic Confinement System," Ph.D. Thesis, MIT, June 1990.

2/5/33

URC97045

RADIATION TOLERANCE OF INTEGRATED CIRCUITS IN SPACE AND THE TERRESTRIAL ENVIRONMENT

T. N. Fogarty, Z. You, K. Washington F. Brown, T. Nichols
Prairie View A&M University- NASA-Center for Applied Radiation Research

ABSTRACT

Current ULSI (Ultra Large Scale Integration) technology contains submicron features. Fabrication processes for this commercial terrestrial application now include many processes capable of producing radiation induced defects in MOS structures. Therefore, a moderate degree of Total Dose immunity in the devices is desirable. Since the demise of the cold war, space avionics has many NASA-DOD applications that have a substantially relaxed total dose requirement. However, at the higher inclination space station orbit, susceptibility to proton-induced single event effects (SEE) is substantially increased. Future generations of commercial ULSI will operate at a lower voltage, reducing noise margin and lower specific capacitance, lowering critical charge, thus susceptibility to cosmic ray induced single event upsets (SEU). A new model is presented for understanding of multiple bit flips in particle accelerator SEU testing. Considering the above, a growing convergence exists between commercial ULSI for terrestrial applications and space avionics.

TOTAL DOSE AND DOSE RATE EFFECTS IN CMOS

In order to obtain the deep submicron feature size of future generation ULSI, an increasing number of fabrication process steps are now capable of producing radiation induced defects which shift device parameters. These include: sputter deposition of alloy conductors to prevent metallic spiking in shallow junction CMOS source and drain applications; plasma deposition of refractory metal vias and inter level dielectrics; plasma and reactive ion etching to pattern fine feature conductors and vias; deep UV, X-ray and electron beam photolithography processes. The fabrication process may include some annealing steps to reduce radiation damage from the above steps and displacement damage from ion implantation. A moderate level of Radiation Total Dose Tolerance is desired. [1,2,3,4]

Since the demise of the cold war, a substantial number of NASA-DOD space avionics applications exist with the Total Dose requirement reduced to moderate levels.

OBSERVATIONS AND HYPOTHESIS ON DEFECTS At The SILICON SILICON DIOXIDE INTERFACE

Equation 1 defines the threshold voltage for a nmos device. (figure1) [3]

$$V_{th} = \phi_{ms} - \frac{Q_{ss}}{C_{ox}} + \frac{Q_s}{C_{ox}} + 2\phi_f \qquad (1)$$

The surface state charge Q_{ss} which is defined by the following equation 2.

$$Q_{ss} = Q_{fix} + Q_{ot} + Q_{it} + Q_b \qquad (2)$$

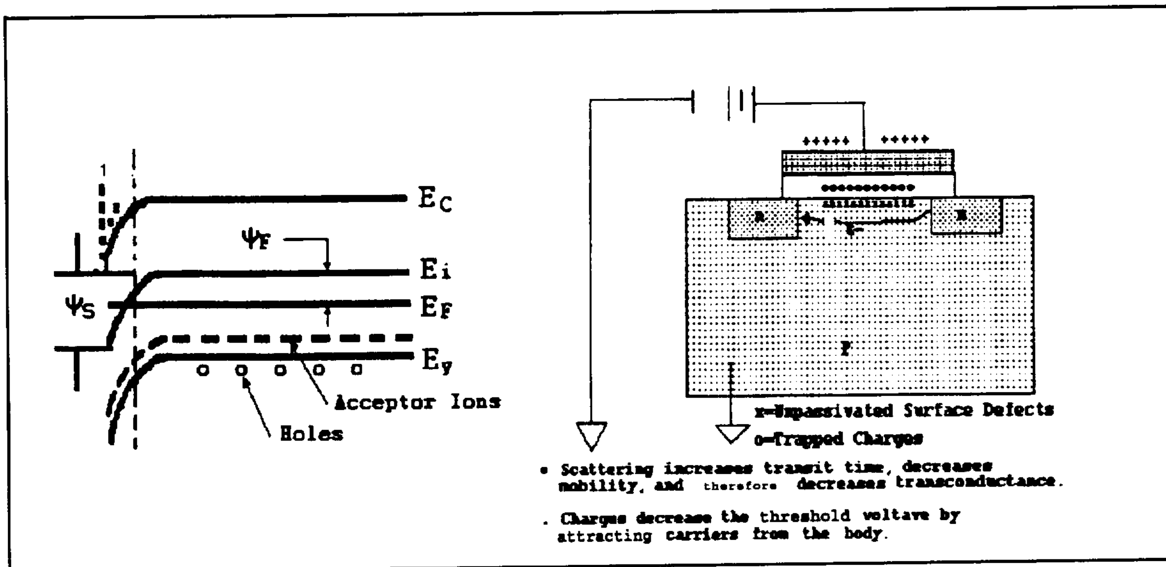


Figure 1. NMOS Transistor

where:

Q_{fix} stands for Fixed positive Charge where small island clusters of silicon exist at the Si-SiO_2 interface.

Q_{ot} refers to Oxide Trapped positive Charge which is an oxygen vacancy in the Silicon Dioxide.

Q_{it} stands for Interface State Trapped Amphoteric Charges which are Silicon dangling bonds.

Q_{bt} are Border Traps. Border Traps. Like Interface States they are in communication with the substrate but are slower traps.

We have been examining radiation and process induced defects at thesilicon/silicon-dioxide interface in the Rad-Hard MOS system for some time. In the previous data, MOS devices were exposed to gamma and proton irradiation at equivalent dose rates. Various electrical parameters were measured before, during andafter the irradiation. The following were observed from these measurements: [5,6,7,8]

1. The average capture cross section of interface states was reduced with 1 Mrad radiation of 1 MeV protons. It was hypothesized that new interface states were created during radiation due to the H^{+} nature of the protons. This was negated because similar results were obtained using gamma radiation, (Fig. 2A, B, &D)

2. For radiation of NMOS devices under inversion bias, acceptor states in the upper half band gap are filled. If the bias is removedafter annealing, electron transport will remove the charge but not the states so that the effect is reversible with bias charge. It was hypothesized that the primary reason for the average capture cross section reduction is the creation of more electron trapping interface states (assuming that electrons have a smaller capture cross section). This was negated because under inversion bias, bothPMOS and NMOS devices exhibit similar capture cross section reductions. (Fig. 2C)

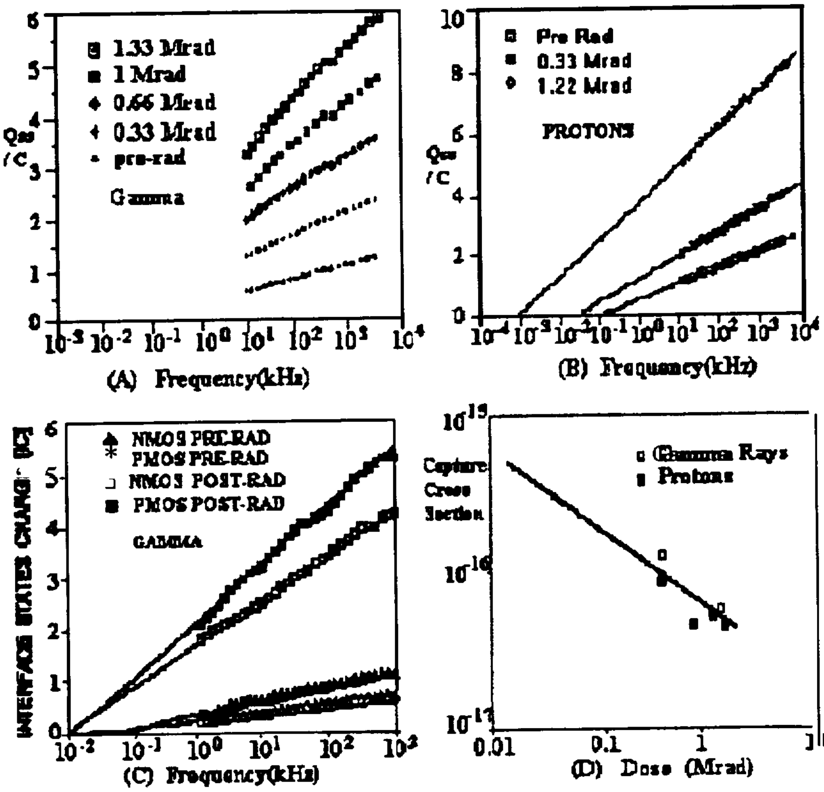


Fig.(3) Interface State Capture Cross Reduction with Radiation

3. The density of interface states increased at approximately the same rate with total dose for both proton and gamma radiation. This implies that the change in capture cross section may depend on two separate species (possibly theP_{b0} and P_{b1} states). It is not clear however, the exact nature of these states. Electron Spin Resonance experiments with P. Lenahan (Penn State) may resolve this issue.

4. Radiation at liquid nitrogen temperature does not create interface states. However, again we find that the capture cross section of the interface states is reduced thus a structural charge in the interface state dangling bond is again suggested.

HYDROGEN EFFECTS

A common practice in commercial CMOS fabrication is post-metallization annealing at 450°- 500°C in a hydrogen atmosphere. This hydrogen is thought to passivate the Silicon dangling bond thus lowering the interface state density. However, with subsequent radiation, depassivation occurs. Therefore, RAD HARD parts are fabricated with silicon dioxide with as little water and hydrogen introduction as possible. Jaccodine has suggested fluorine incorporation in the oxide which increases the oxidation rate and replaces the hydrogen at the dangling bond providing a more stable healing of the interface state and a degree of immunity to subsequent radiation [3,4,9,10]. Large NMOS MOSIS chips had similar results to previously tested RAD Hard chips; but initial capture cross section was larger indicating a different interface structure. [7]

SINGLE EVENTS **EFFECTS**

Single Event Effects [7,8,9] are usually simulated on earth by Particle Beam Accelerator. The effective Linear Energy Transfer $LET_{eff}=LET/\cos(\theta)$ which θ is the particle angle of **arrival:**

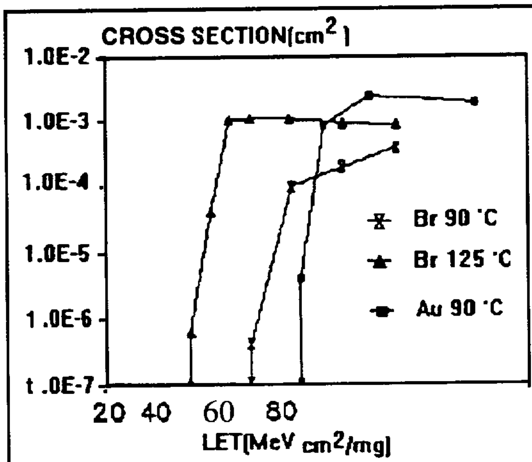


Fig. (3A) Comparison LET Threshold for Bromine and Gold SEE Test

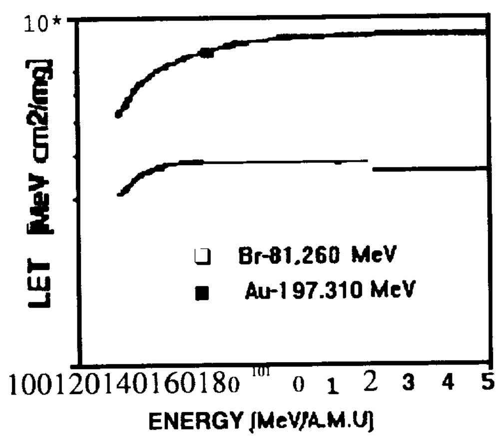


Fig. (3B) LET[Si] from ZIEGLER

The concept of effective LET has been questioned. Therefore, RAD HARD SRAM devices were exposed to bromine and gold ions at BNL Van De Graaff Accelerator. A difference was observed. (Fig. 3). This maybe caused by energy lossin the passivation layer. The LET versus energy curves for bromine and gold ions in silicon are plotted in Fig. 3B. A small energy loss shows in the 260 MeV ^{81}Br ions gradually approaching a maximum, thus the LET remains relatively constant. On the other hand, 310MeV ^{197}Au ions have already passed the maximum and are in the region of steep decay, thus a significant LET reduction is expected. [4,5] We have shown that prior total dose radiation increases SEE[5]. **C. Lage et.al** of Motorola has predicted cosmic **ray** induced SEE on earth with future generation **ULSI. Also IBM has devoted an entire issue of their Journal of Research and Development (January 1996) to terrestrial radiation effects.**

Recently, Woodruff and Rudeck [10] and Sexton have used DAVINCI (TMA'S 3 Dimensional Version of PICES 2B, a Poisson's equation and continuity equation solver for device geometry) to show that LET effective and LET critical depend on the sense of ø. (Arrival Angle of Cosmic Ray or Particle Beam at Device Under Test, eg. +45° caused SEU while -45° did not). This depends on the exposure and charge collection at critical nodes of the SRAM. In mirrored cell lay outs only one half the memory cells may be critically tested for SEE, The same 16K SRAM were tested at the TAMU Cyclotron with following results:

Particles	Angle θ(degree)	LET_{eff}(MeV cm^2/g)	SEU
^{84}Kr	0	28.3	0
^{84}Kr	±45	39.9	0
^{93}Nb	0	34.5	0

Therefore, it appears that critical LET or LET threshold are lower on the Cyclotron than *on* the BNL Van De Graaff. This may be partially due to much higher energy in relation to the Bragg Peak of the Cyclotron. Also at LET_{eff} 39.9 with $\theta = \pm 45$ there was zero SEU. This may be partially explained by the Woodruff prediction in mirrored cells.

MULTIPLE SEU

In multiple SEU every even SEU induces the cell to its correct values. So the number of errors increased in the time period dt is:

$$dN_e = \frac{(N - N_e(t))}{N} R_{seu} P_r f dt - \frac{N_e(t)}{N} R_{seu} P_r f dt$$

Figure 4. la is the result of a numerical solution of the above differential equation by using a small computer program. The result shows that no matter how long you expose the SRAM to the radiation, the maximum number of errors is half of the total SRAM cells. Figure 4. lb is the test results using Ar and Kr on 64K SRAM. From the results we can see that Kr has a much higher SEU susceptibility because of higher LET, so this model is consistent with the model using capture cross section vs. LET. [15]

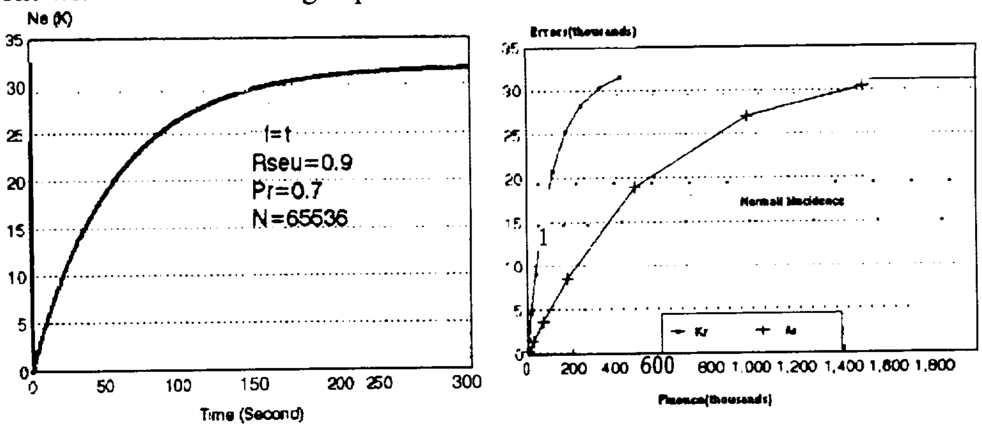


Fig 4 Numerical Solution and Testing Results

APPROACHES TO A MODERATE RAD TOLERANT ULSI TECHNOLOGY

Future approaches to a convergent RAD tolerant commercial ULSI for space avionics will include: [4,5,9, 10,12,13]

-Fluorine stabilization of gate field oxides.
-Increased critical charge by high dielectric constant gate insulator,
-Innovative circuits increase RAD tolerance of commercial ULSI.
-Horizontal and vertical ULSI device design to reduce Latch-Up.
-Awareness that prior total dose radiation may lower LET threshold.
-Resolution of SEE test anomalies.

REFERENCES

1. S. Wolf and R. N. Tauber ed. "Silicon Processing for the VLSI Era," Vol. I-Process Technology, Lattice Press, California 1986
2. S. Krishnan, S. Aur, G. Wilhite and R. Rajgopal, "High Density Plasma Etch Induced Damage to Thin Gate Oxide", IDEM 95-315, page 13.2.1

3. T. P. Ma and P. V. Dressendorfer. "Ionizing Radiation Effects in MOS Devices and Circuits," John Wiley and Sons, New York, 1989.
4. T. N. Fogarty "Process and Radiation Induced Defects in Electronic Materials and Devices" Invited paper ACerS May '95; published in "Crystal Growth of Novel Electronic Materials", Ceramic Transactions, Volume 60.
5. V. Zajic, T. N. Fogarty, R. Kohler &E. G. Stassinopoulos, "Single Event Upset and Total Dose Radiation Effects on Rad Hard SRAMS", Journal of Electronic Materials, Vol. 19, No. 7, 1990.
6. T. A. Thomas, T. N. Fogarty, D. Lin, K. J. Kloesel, "Detection of Process and Radiation Induced Defects in MOS Devices"; PVAMU-NASA/JSC Conference on Materials, Devices, Circuits and Biosystems, November 22 & 23, 1993, Page 22 ff.
7. D. Lin, L. Trombetta, T. N. Fogarty, "Formation of Interface Traps in MOSFETS from 77K to 323 K", Radiation's Studies Conference, NASA/Center for Applied Radiation Research, Prairie View A&M University, page 1.
8. Len Trombetta, T. N. Fogarty, "Slow States, Anomalous Positive Charge, and Border Traps in MOS Devices", Radiation Studies Conference, NASA/Center for Applied Radiation Research, Prairie View A&M University, page 4.
9. V.S. Kim, C.H. Walowodieck, R.J. Jaccodine, F. Stevie and P. Kahoro. "Effects Fluorine Additions to the Oxidation of Silicon", J. Electrochem. Sot., Vol. 137, No. 7, July 1990.
10. T.P.Ma, "Metal Oxide Semiconductor Gate Oxide Reliability and the Role of Fluorine", J. Vac.Sci.Tech. A, Vol 10, Pg.705, 1992.
11. T. N. Fogarty, J. O. Attia, A. A. Kumar, T. S. Tang and J. S. Linder, "Modeling and Experimental Verification of Single Event Upsets" Selected Topics in Robotics for Space Exploration (Ed. R. C. Montgomery) NASA Conf. Pub. 10131, NASA LaRC, Dec. 1993.
12. C. Lage, D. Burnett, Thomas McNelly, K. Baker, A. Bormann, D. Dreier, V. Soorholtz, "Soft Error Rate and Stored Charge Requirements in Advanced High-Density SRAMs," IEDM 93-821, page 33.4. I
13. R. L. Woodruff and P. J. Rudeck, "Three Dimensional Numerical Simulation of Single Event Upset of an SRAM Cell," IEEE TRANS. NUC.SCI., Vol. 40, No. 6, Dec. 1993.
14. F. W. Sexton, K. M. Horn, B. L. Doyle, ET.AL., "Relationship Between IBICC Imaging and SEU in CMOS IC's," IEEE TRANS.NUC.Sci., Vol. 40, No. 6, Dec. 1993
15. Z. You, "SRAM Single Event Upset Dependence on Accelerator Type, Beam Arrival Angle (+/-), and Ion Species" MS Thesis, PVAMU/CEA, August 1995

URC97046

Creation of *Quantity*: An Object-Oriented Reusable Class

Monica Gayle Funston & Robert Valerio
Graduate Students, University of New Mexico, Department of Civil Engineering

Walter Gerstle, Ph.D.
Associate Professor, University of New Mexico, Department of Civil Engineering

Abstract

More every day, computers are handling many of the repetitive tasks for the engineer. One of them is quantity manipulation. This paper discusses the necessity of the class *Quantity*. *Quantity* has the ability to maintain the full meaning of a quantity and to handle operations such as unit conversion and coordinate system transformation. This paper also discusses how object-oriented design enhances *Quantity* with the valuable characteristics of generality and reusability.

Introduction

Webster's dictionary (Anonymous, 1956) defines quantity as

1. that property of anything which can be determined by measurement. . . .
7. in mathematics,
   a) a thing that has the property of being measurable in dimensions, amounts, etc., or in extensions of these which can be expressed in numbers or symbols,
   b) a number or symbol used to express a mathematical quantity.

The ability to quantify the world in which we live and work is the essence of engineering and science. Engineering and science demand a multitude of ways to manipulate numbers and values, and an engineer needs a mechanism with which to deal with quantities. The engineer must have the ability to use values of any type. However, the engineer does not work with values alone, he also uses a standard of measurement, called a unit, to bring meaning to the numbers. (Sixty degrees Celsius has a much different meaning than sixty seconds,) Furthermore, a quantity sometimes needs more than just a magnitude and a unit of measurement. A value, even if having specified units, needs a frame of reference in the form of a coordinate system. For example, a velocity, 550 mph, is given more meaning if you make it into a vector by giving it direction such as, "70° East of North", Furthermore, even if we know that the quantity is "550 mph, 70° East of North," it is helpful to give the quantity additional meaning with an annotation such as "velocity of Boeing 727 between Albuquerque and Dallas." An annotation is helpful to keep track of the use and implementation of the quantity in the engineering problem. Oftentimes, an engineer also finds it necessary to use different systems of units, ft in the US Standard and kg in the S.I. system, or to switch from a larger to a smaller unit; for example, to change m to cm or to change yd to ft. An inordinate amount of valuable time is spent converting values to the right magnitude to reflect the unit change and errors often arise in proportion to the number of conversions needed.

Computers are handling more of the repetitive tasks for the engineer everyday. When the engineer desires to work with a large database of values and units of different types and/or systems, it is logical for her to turn to a computer to do much of the quantity manipulation.

The need for a software component to manipulate quantities materialized during the development of a computational mechanics toolkit at the University of New Mexico. The

toolkit's purpose is to model, assemble, and analyze engineering boundary/initial value problems. A search was made to find a reusable software component with capabilities to:

store data as values with respect to units and coordinate systems.

- perform arithmetic operations between quantities.
- perform unit conversions.
- perform coordinate transformations.

We searched for software with these capabilities. Some of these characteristics were found in several commercial programs such as : I-DEAS by SDRC (I-DEAS, 1995), Mathcad by MathSoft (Mathcad, 1992), and MultiFrame3D by Graphic Magic (MultiFrame3D, 1990). However, these applications are proprietary and, therefore, not reusable, at least not in academia. We also sent a query on a listserver called feusers-request@mailbase.ac.uk and received some responses, but no shareware that fulfilled the criteria. This lack of a reusable component with these capabilities prompted the project to design the class *Quantity*.

One approach to modern software design is the object oriented paradigm. Due to the inherent characteristics of reusability and maintainability, object oriented design was the chosen technique with which to design *Quantity*. Especially in the academic arena where the trend toward an object oriented approach to software design and the use of C++ to program is growing steadily, the existence of such a fundamental thing as *Quantity is* essential.

Design of *Quantity*

Only when a class is understood thoroughly can it be implemented, so the first step to the creation of *Quantity* was the formulation of a set of user and developer requirements and of a working vocabulary with clear, concise definitions (please see glossary at the end of paper). With the requirements and the vocabulary as a guide, a class structure was built along with the assignment of corresponding responsibilities for each class (Wirfs-Brock et al., 1990). With the needs of both the application user and the developer held paramount, *Quantity* has gone through many design cycles. Every iteration in the process spiraled closer toward both a clear and concise definition and toward a more useful design of *Quantity*.

For the application user, *Quantity* needs to be an object created to simulate its engineering and scientific function as closely as possible and is comprised of a value, a unit, a coordinate system, and an annotation. For the developer the main design goals for *Quantity are* to be:

- easily reused by other programs.
- as general as possible.
- reasonably efficient.

Quantity was designed as a base class, meaning that every type of *Quantity* implemented has common characteristics that it inherits from the base class. The *Quantity* base class contains:

- a UnitSymbol
- a Coordinates ymbol
- an Annotation

(The *Value* is a component of all classes derived from *Quantify.)* The UnitSymbol defines what standard of measurement is used to define the quantity and the CoordinateSymbol distinguishes which coordinate system is used to define the quantity. The Annotation explains how the quantity is used in the engineering problem.

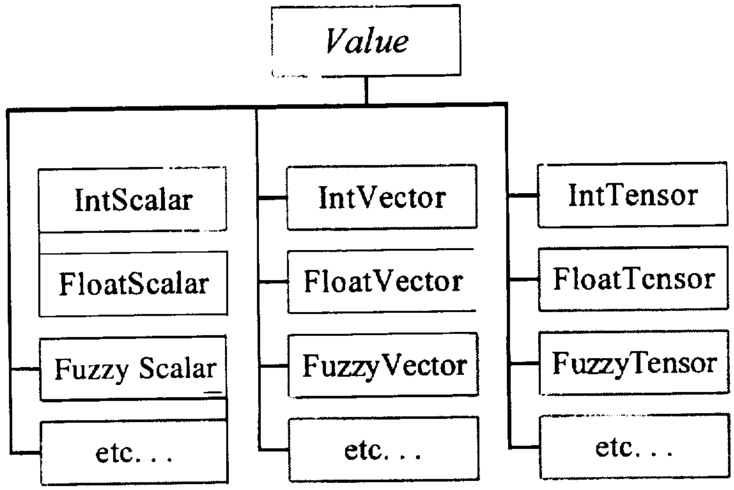


Figure 1. Value Hierarchy Chart

Each derived *Quantity* also contains a numeric component, the *Value*. There are two categories with which to define the type of *Value*. The first category specifies whether the value is a float, integer, or fuzzy number, etc., and the second category specifies whether the value is a scalar, vector, or a tensor. *Value is* the base class, but the developer will use only its derived classes to define the type of value the Quantity will contain. To define length, the developer may choose to use the FloatScalar value type such as, "4.59" or a IntScalar such as "5".

Because units are an integral part of a *Quantity,* the ISO 1000: International Standard (ISO, 1992) is an invaluable basis to use in the creation of *Quantity*. The ISO Standard states that all units are derived from a set of base quantities (not to be confused with the base class *Quantity*) which are comprised of the seven physical dimensions (length, mass, time, temperature, luminous intensity, amount of substance, and electric current). The fundamental definition of a unit is a set of BaseQuantityExponents applied to each of the base quantities. Depending on which system of units is in use, each base quantity has associated with it a BaseUnit to which the exponents can be applied. A *UnitDictionary* is needed to hold all the information needed to properly define a unit, as shown in Table 1.

UnitDictionaryEntry

Unit Symbol	**Conversion Factor**	**Base Quantity Exponents ***						
		L *m*	*M* *kg*	*t* *sec*	*T* *C“*	*LI* *c*	*AS* *mol*	*EC* *ohm*
ft	*0.302*	1	0	0	0	0	0	0
N	1	1	1	- 2	0	0	0	0
MPa	1000000	1	1	- 2	0	0	0	0
m	1	1	0	0	0	0	0	0
m^2 / s^2	1	2	0	- 2	0	0	0	0

*L= Length, M= Mass, t= Time, T= Temperature, LI= Luminous Intensity, AS= Amount of Substance, EC= Electric Current

Table 1. UnitDictionary

The *UnitDictionary* stores a ConversionFactor to the default unit system for each UnitDictionaryEntry. The default unit system adopted in our implementation is the S.I. unit

system. With all of these elements: UnitSymbol, ConversionFactor, and BaseQuantityExponents any UnitDictionaryEntry can be completely defined. The addition of new UnitDictionaryEntries to the *UnitDictionary* is simple, because of the consistent way of defining units.

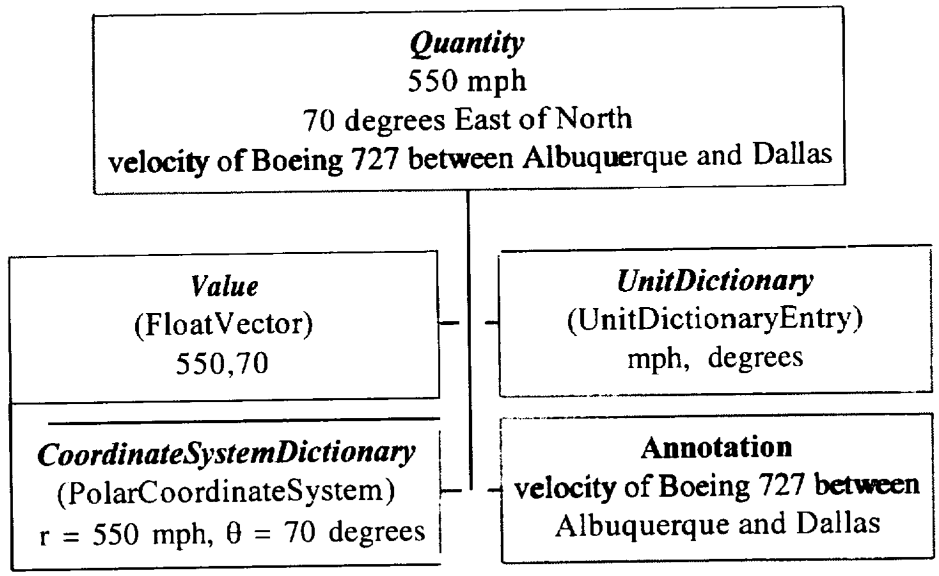


Figure 2. Components of *Quantity*

By collaborating with the *UnitDictionary*, *Quantity* solves the many intricacies involved in performing *Value* conversions with respect to Units. The implementation of value conversion due to a unit change occurs in two steps. First, *Quantity* queries the *UnitDictionary* to *see* if the conversion is legal, meaning that the BaseQuantityExponents of both Units involved are the same. For example, feet can be converted to meters or yards, but not kilograms. Then if the Unit can be converted, *Quantity* queries the *UnitDictionary* for the Conversion Factor and the value conversion takes place in *Quantity*.

A *CoordinateSystemDictionary is also* devised. Just as the *UnitDictionary* has a default system of base units, the *CoordinateSystemDictionary* adopts the rectangular Cartesian coordinate system as its default coordinate system. The *CoordinateSystemDictionary* also contains definitions of different systems and the rules for transformation from one system to another. For example, in the Cartesian system, the coordinates (x, y, z) specify a point and in the cylindrical coordinate system, the radial, angular, and axial coordinates (r,θ,z) specify a point.

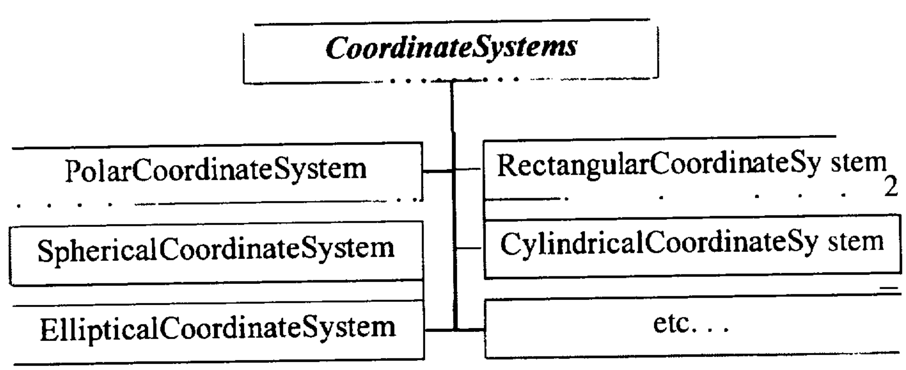


Figure 3. CoordinateSystem Hierarchy Chart

An Example for the Need of Quantity

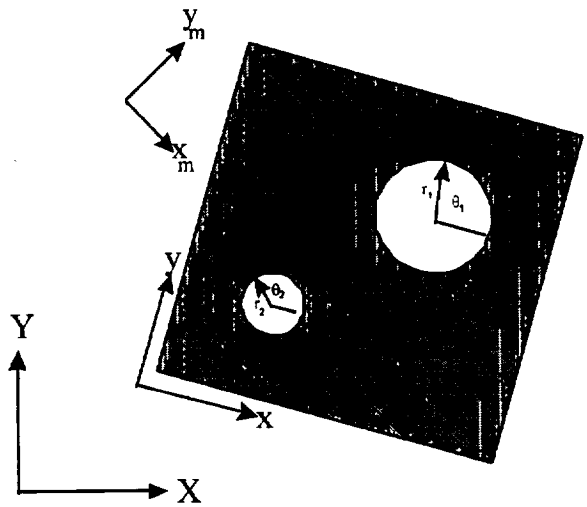


Figure 4. Plate with Bolt Holes.

Figure *4* is a plate with bolt holes drilled into it and is an excellent example of when it is useful to have *Quantity* and its ability to use different coordinate systems and unit systems. For example, each piece of geometry may be conveniently positioned in its own separate coordinate system, In addition, stresses and strains can be conveniently considered in their own geometrical coordinate systems (x,y) or in the material coordinate system (y_m, x_m). To illustrate the usefulness in using different unit systems, it maybe easier to designate the bolt holes on the plate in metric units and the plate dimensions in U.S. Custom units. The stresses and strains can also be specified in either unit system.

Conclusions

Quantify provides the user with the ability to easily combine any unit configurations and coordinate systems desired. *Quantity* is designed to be compatible with any object-oriented program needing to keep track of and store objects with values, and associated units, coordinate systems, and annotations.

Quantity has been implemented and tested in a stand-alone test program. Although it does not yet have all the criteria it had been designed to possess, namely the *CoordinateSystem-Dictionary*, *Quantity* has been able to successfully test the other characteristics.

With the help of the object-oriented approach, *Quantity* has developed into a well thought out piece of software. With *Quantity's* modular design, it can be maintained and modified for use in future applications needing *Quantity*'s capabilities. Its versatility also allows the user to use it as a stand-alone application. *Quantity is* presently being tested in the computational mechanics toolkit being developed at the University of New Mexico.

Glossary of Terms for *"Quantity"*

Quantity - the physical meaning of a number as a *Value* with associated Unit and Annotations

Value - the numeric part of *Quantity*

Unit - symbols/names that represent a standard of measurement of *Quantity* (m, see, m/s^2, ***N***, meter etc.)

Annotation - additional information to a *Quantity* to help application user with identification, notes, modification date, etc. (user notes attached to a *Quantity*)

UnitDictionary - a list of UnitDictionaryEntries

UnitDictionaryEntry - contains definitions of each Unit in terms of the default unit system (S.I.) using Base Quantities and Base Units:

a. Unit
b. BaseQuantityExponents
c. ConversionFactor

Base Quantity - one of the seven physical dimensions per ISO 1000 pg. 2 (length, mass, time, temperature, etc.)

BaseUnit - the unit representing one of the seven Base Quantities using the S.I. system of units (m, kg, see, K, candela, mole, and ampere)

Base Unit Exponent - power the Base Units are raised to define a Unit

ConversionFactor - the factor by which to multiply the *Value* part of *Quantity* from a given Unit to the Base Units

CoordinateSystemDictionary - a list of CoordinateSystemEntries

CoordinateSystemEntry - contains the definitive components needed to fully represent a *Quantity* in a specific coordinate system (Polar - [r,θ], Rectangular Cartesian - [x,y])

References

1. Anonymous, "Webster's Dictionary of the American Language College Edition", World Publishing Company, New York and Cleveland, 1956.
2. Wirfs-Brock, Rebecca, Brian Wilkerson, Lauren Weiner, "Design of Object Oriented Software", Prentice Hall, Englewood Cliffs, NJ, 1990.
3. "I-DEAS Version 2.1", Structural Dynamics Research Corporation (SDRC), Milford, OH, 1995.
4. "Mathcad Version 3.1", MathSoft Inc., Cambridge, MA, 1992.
5. "MultiFrame3D Version 1.0", Computer Aided Design for the Structural Engineer, Graphic Magic, Inc., Santa Cruz, CA, 1990.
6. "ISO 1000: International Standard", Third Edition 1992-11-01, Reference Number ISO 1000:1992(E).

N 1 1 5 /
/ 5 3

URC97047

A LOW COST RAD-TOLERANT STANDARD CELL LIBRARY†

Jody W. Gambles   Gary K. Maki
NASA Institute of Advanced Microelectronics
Microelectronics Research Center
University of New Mexico
Albuquerque, New Mexico 87131 USA
Tel: 505-272-7040 FAX: 505-272-7041
jgambles@mrc.unm.edu
gmaki@mrc.unm.edu

Abstract

This paper describes circuit design techniques developed at the NASA Institute of Advanced Microelectronics that have been shown to protect CMOS circuits from the deleterious effects of the natural space radiation environment. The IAμE is leading a program to incorporate e these radiation-tolerance providing design techniques into a commercial standard cell library that will be used in conjunction with available Electronic Design Automation tools to produce space flight qualified microelectronics fabricated at modern commercial CMOS foundries.

1 Introduction

Microelectronics used in space systems are subjected to the deleterious effects of the natural radiation environment found outside of the protection of the earth's atmosphere. During the 1970s and 1980s the United States Departments of Defense (DoD) and Energy (DoE) sponsored the development of radiation-hardened semiconductor processes. Government space agencies and the commercial satellite industry have been able to utilize many of these rad-hard components to increase system complexity and reliability while reducing size, weight, and power requirements for space-borne platforms. In the post-Cold War world the DoD and DoE push to cent inue to advance the rad-hard processes has waned. As a result, the performance capabilities of the available rad-hard components have lagged behind those that are manufactured using the latest commercial technologies. Space craft designers are facing an ever widening performance gap between available rad-hard and commercial devices [1]. Compounding the problems faced by the satellite industry, rad-hard components are becoming harder to get. At least six sources of rad-hard parts have exited this market in the last five years, leaving only two domestic suppliers. This paper describes design techniques which produce rad-tolerant CMOS circuits, and outlines a path currently being pursued to provide this technology to designers of Application Specific Integrated Circuits (ASICs).

2 Ionizing Radiation Effects in MOS Microelectronics

Ionizing radiation may be defined as exposure to charged particles that possess enough energy to break atomic bonds and crest e electron/hole pairs in the absorbing material. Such particles may include protons, electrons, atomic ions, and photons with energies greater than the material bandgap. There are two primary categories of ionizing radiation effects in microelectronics; total ionizing dose (TID) effects and transient effects [2]. TID effects are a function of ionizing radiation accumulation over months or even years, which can lead to performance degradation and functional failure. Transient radiation effects are primarily the result of phot ocurrents generated as energetic particles pass through the circuit.

† This work is being supported by NASA under the Institute of Advanced Microelectronics grant NAGW-3293.

2.1 Total Ionizing Dose Effects

As ionizing particles pass through MOS devices, generating electron/hole pairs, charges can be trapped in the gate and field oxides and interface states are increased. Mobile electrons quickly transport through the oxide, but holes have a very low effective mobility in SiO_2 and are easily trapped. The trapped positive oxide charge shifts transistor threshold voltages in a negative direction. An increase in interface states shifts thresholds in the positive direction for n–channel devices and in a negative direction for p–channel devices. Generally, the trapped oxide charge shift is greater than the interface states shift and the magnitude of the NMOS V_t decreases while the magnitude of the PMOS V_t increases. The radiation–induced interface states also reduce the channel mobility, which decreases channel conductance reducing the transistor drive. Over time, the threshold voltages may shift to the point where the n–channel transistors cannot be turned off and the drive capability of the p–channel transistors is not sufficient for the chip to continue operating at the system clock rate, causing it to fail, In addition to the drawn transistors, threshold shifts also occur for parasitic MOS elements. As the parasitic n–channel transistors thresholds decrease, channels begin to form and leakage currents flow around the edges of the drawn n–channel gate regions, from drain to source, between drain/source regions of adjacent n–channel transistors, and from n–channel drain/source regions to the n–well/n–substrate. Leakage currents may cause parametric failures to occur before functional failures.

While the actual dose that a satellite receives is highly dependent on the orbit, satellites in low earth orbit can be expected to receive a TID exposure of less than 10K rads(Si) during missions of up to 20 years. For a satellite in geosynchronous orbit the TID can be expected to reach 10 OK rads(Si) after 10 years on orbit. The most severe ionizing radiation orbits are 1/2 geosynchronous, which can reach a 1M rad(Si) dose after 8 years on orbit [1]. The radiation hardness of a MOS process is a function of the rate at which oxide–t rapped charge and interface traps build up as the radiation dose increases. Scaling of commercial processes has naturally reduced the volume of the gate oxide and thus reduced drawn transistor threshold shifts, leaving leakage currents as the dominant TID effect. Some commercial processes have been shown to produce parts that exhibit TID hardness in the 100's of KRads [3, 4].

2.2 Ionizing Radiation Transient Effects

Single Event Effects (SEE) are produced in the natural space environment by galactic cosmic rays, solar enhanced particles and energetic protons and neutrons [2]. The passage of a single high–energy particle through a MOS device can create a high–density track of electron/hole pairs which results in charge collection in a localized region of the circuit. SEES are commonly divided into two categories, Single Event Latchup and Single Event Upset.

In complementary MOS (CMOS) devices containing both n–channel and p–channel devices on a silicon substrate, parasitic bipolar p-n-p-n devices exist, forming a silicon–cent rolled rectifier (S CR) structure, which under normal conditions is in its "off" (i.e. high–impedance) state. If a SEE injected photocurrent produces sufficient bias to turn on one of the parasitic base-emitter junctions, the SCR can be triggered, producing a low–impedance path between the power supply and ground rails. If the product of the effective current gain (β product] of the parasitic p-n-p and n-p-n devices is greater than unity, then a regenerative condition exists and a self–sustaining SCR high current mode is entered after the triggering event [5]. This condition is know as Single Event Lat chup (S EL) and can cause destructive failure. The SEL phenomenon is similar to the Electro–Static Discharge induced latch-up protected against in typical CMOS 1/0 structures, however in an ionizing radiation environment, a particle can strike anywhere in the circuit so merely protecting the 1/0 circuitry is not sufficient.

A Single Event Upset (SEU) occurs when the charge transferred as a result of the generated photo currents is of sufficient magnitude to alter the logic state of a susceptible node. An upset node may further cause the alteration of the contents of circuit memory elements or alter the operation of the circuit in such a way to cause an error in the logic function.

3 Ionizing Radiation Effect Immunity By Design

One spacecraft design approach is to address SEES at the system level while using commercial *off the shelf* (COTS) parts. Limiting the supply current to a device can save it from latchup destruction, but requires a power down and reset cycle whenever a SEL occurs. Logic malfunctions due to SEUs can be detected and corrected through system level redundancy. However, this strategy can be quite costly and still leaves unanswered the classic question, "who checks the checker?"? It has been demonstrated that circuit and layout design techniques can make it possible to provide a high degree of SEL and SEU immunity using commercial CMOS processes. In general, these techniques do increase cell area, decrease speed, and/or increase power consumption. An optimal solution should minimize these costs.

3.1 SEL Immunity

The techniques used to prevent latchup in CMOS devices involve degrading the β product of the parasitic n-p-n and p-n-p transistors and/or limiting the applied base bias [5]. Approaches to β product reduction include minority carrier lifetime degradation in the parasitic base (i.e. substrate and well) regions, accomplished by gold doping [6] or neutron irradiation [7]. Insuring some minimum spacing between source/drain regions in the substrate and the well edges decreases the β of the lateral parasitic by insuring a wide effective base region [8]. The base bias is reduced by lowering the effective base-emitter resistance in the parasitic SCR structure. Low–resistance connections from the substrate and well to the power and ground rails also reduce the base bias current by providing for capture and shunting away of injected minority carriers before they reach the parasitic base. Methods for reducing the substrate and well resistance and increasing charge carrier capture include the use of a lightly doped epitaxial layer on top of a heavily doped substrate [9] and the use of p^+ guard rings around the n–channel transistors and n^+ guard rings around the p–channel transistors [10]. It has been shown that latchup can occur in circuits fabricated using an epi layer process [11], and that the epi layer must further be "thin" in order to prevent latchup.

The minority carrier lifetime degradation and thin epi–layer solutions belong to the "technology hardening" class of solutions. These approaches rely on specifying and/or controlling some aspect(s) of the fabrication process and are not generally considered to be "commercial" CMOS. Guard rings are produced during the normal source/drain mask steps and require no special processing. The guard ring method has been shown through heavy ion testing using the Twin Tandem Van de Graaff accelerator at the Brookhaven National Laboratories (BNL) Single Event Upset Test Facility (SEUTF) to prevent SEL at LET levels ranging from 3.4 $MeV{\cdot}cm^2/mg$ up to at least 120 $MeV{\cdot}cm^2/mg$. These results have been obtained using multiple test chips fabricated through MOSIS in Hewlett Packard's $1\mu m$ double metal CMOS (CMOS34) process, and Hewlett Packard's $0.8\mu m$ triple metal CMOS (CMOS26b) process, as well as a 1.2 million transistor radio astronomy correlator chip implemented in the CMOS26b process, a 100,000 transistor Reed–Solomon error correcting code (ECC) encoder and a 200,000 transistor Reed–Solomon ECC encoder/decoder, both fabricated in American Microsystems Inc. (AMI) triple metal $1.0\mu m$ process (CYC) [12, 13]. The cell area cost of including guard rings scales with reduced feature size while the minimum spacing approach does not scale. A comparison of the results of Moss et. al. [8] with the IAμE results at BNL show that the guard ring met hod cost is lower for sub–micron processes.

3.2 SEU Immunity

Multiple strategies have been applied to harden microelectronic circuits against the effects of SEU. One approach is to reduce the charge collection capability of the material to the point that the circuit will not collect sufficient charge to initiate an upset [2]. This strategy belongs to the 'technology hardening" class. Other circuit design based approaches seek to raise the critical charge required to upset sensitive storage nodes. Finally, redundancy techniques have been applied at the circuit level to recover upsets. The primary goal of SEU hardening through circuit design techniques is to produce SEU–immune circuits using standard CMOS processing, with no additional mask or processing steps, while minimizing cell size, circuit speed costs, and power consumption.

The enhanced critical charge hardening techniques include increasing transistor drive, capacitive hard-

ening, and resistive hardening. A high drive transistor can quickly remove/replace SEU injected charge, shortening the time duration of the disturbance. Large high drive transistors also have increased node capacit ante, which reduce the voltage excursions caused by the SEU injected charge [14]. Increasing the capacit ante of critical nodes to reduce the voltage change due to SEU injected charge is the basic concept behind capacitive hardening of circuits [2]. Resistive hardening involves the use of resistors in the memory element feedback paths, to create, in conjunction with the gate capacitance, a low pass filter to reject the effects of SEU induced transients while passing the longer duration legitimate *signals [15]*.

Power is consumed charging/discharging circuit capacitance each t ime the logic level of a node changes. Increasing circuit capacitance due to high drive transistors or other capacitive hardening methods also *increases* the ac power consumption of the circuit. Designing a cell to reject short duration signals places a constraint on the maximum speed at which the circuit can operate. Under *nominal conditions,* it is possible to design an RC filter to reject SEUs and still allow the circuit to operate at hundreds of megahertz. The resistances required to provide SEU immunity are typically 10 OK to 1M ohm, requiring high resistivit y polysilicon in order to keep the resistor elements small. High resistivit y polysilicon is very sensitive to doping concentration and therefore subject to wide resistance variations across a generally accepted variation in commercial processing parameters. The polysilicon resistance value control problem due to wafer processing is further exacerbated by a very large negative temperature coefficient. The result is that a cell properly designed to reject SEUs at one corner of the process parameter/operating condition design space suffers adverse performance impacts at other design space corners.

SEU hardening by redundant circuit design approaches are baaed on three fundamental concepts:

1. Information storage redundancy maintains a source of uncorrupted data after an SEU.
2. Feedback from the non–corrupted data storage location should cause the corrupted data to recover after a particle strike.
3. The "intelligence" needed in the feedback to cause recovery of the proper location can be derived from the fact that the current induced by a particle hit flows from n-type diffusion to p–type diffusion. If a memory cell is constructed from only p–type transistors then it cannot be upset to a O while storing a 1. A memory cell constructed from only n–type transistors cannot be upset to a 1 while storing a O.

The low power Whitaker cell[16, 17], developed at the IAμE and shown in Figure 1, consists of two loadable storage structures. The lower storage structure is a modified six transistor cell consisting of only n–type devices and the top structure is a modified six transistor cell consisting of only p–type devices. The lower structure stores incorruptible 0s and the top structure stores incorruptible 1s. In order for the feedback mechanism to effect recovery from SEU, transistors M2 and M4 are sized to be weak compared to M3 and M5 while M13 and M15 are sized to be weak compared to M12 and M14. Complementary n–channel devices M16 and M17 are added to disconnect the de-current path in the p-channel section eliminating the static power consumption that otherwise results from the weak 1 level produced at N1(N2) not turning M 13 (M15) completely off. Similarly p–channel devices M6 and M7 are added to disconnect the de-current path in the n–channel section. The other transistors are sized using the normal design considerations for a memory cell to meet the performance required. The cell buffer transistors M8, M9, M 18, and M 19 restore the output voltage levels to the rails, isolate the storage nodes from high capacitance loads, and tri–state the cell output during SEU recovery. By tri–stating during SEU recovery and not driving outputs to upset values, the capacitance on the cell output maintains the correct output voltage levels during recovery. Supplying separate input signals to the n–channel and p–channel sections combined with an inherent feature of the cell that requires both inputs to be the same in order for the value stored in the cell to change eliminates the capture of propagated upsets that are coincident with clock edges.

The number of transistors required for the SEU–hardened data latch shown in Figure 1 make it impractical for large static memory arrays. However, the design can easily be used to create SE Unhardened master-slave D–flip flops to design finite state machine controllers and other data path elements.

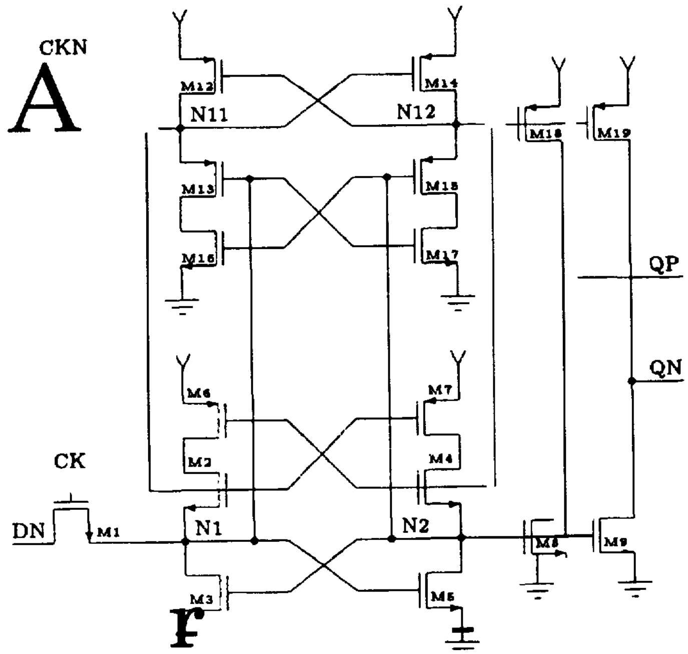


Figure 1: Buffered Low Power Whitaker Cell.

3.3 Rad–Tolerant VLSI Processors

In addition to several test chips, three special purpose rad–tolerant VLSI processors have been developed at the NASA Institute for Advanced Microelectronics utilizing guard rings for SEL immunity and using the low power Whitaker cell for SEU immunity. The rad–tolerant special purpose processors include:

- Error–correcting code (ECC) encoder that supports the Reed–Solomon (RS16) coding specified in the Consultative Committee for Space Data Systems (CCSDS) recommendation for Telemetry Channel Coding.
- Programmable Reed–Solomon ECC encoder/decoder (EDAC) [13]. This chip has been designed into solid–state recorders in support of EOS–AM, LandSat 7, and the Hubble '97 Upgrade Package.
- A 1024 channel autocorrelator chip used in the Naval Research Laboratories (NRL) Orbiting High Frequency Radio Interference Monitor (OHFRIM) experiment [18].

4 Future Work

Designing full–custom rad–tolerant VLSI processors to be fabricated at commercial foundries has been accomplished. However, for this technology to be truly valuable it must be made readily available to a wide range of space system designers. Under sponsorship of NASA, the Institute for Advanced Microelectronics is leading a consortium including academic, industrial, and government partners, to create, qualify, and make available a radiation–tolerant standard cell library utilizing the design techniques described in this paper that will be able to be targeted from a wide range of Electronic Design Automation (EDA) tool environments and fabricated at commercial foundries, The program includes participation by TRW, The Aerospace Corporation, NASA Goddard Space Flight Center and Johnson Space Center, and Aspec Technology, Inc.

References

[1] A. M. Miscione. Radiation-Tolerant Library Development. TRW briefing to Sam Venneri at NASA Headquarters, January 23, 1996.

[2] T. P. Ma and P. V. Dressendorfer. *"Ionizing* Radiation *Effects In MOS* Devices *And Circuits"*. John Wiley & Sons, 1989.

[3] J. R. Chavez, W. T. Kemp, D. A. Bruner, R. D. Bellem, D. R. Alexander, and D. G. Mavis. "Development of Radiation Hardness Evaluation Concepts for Commercial Foundries". In *Digest of Papers: 1996 Government Microcircuit Applications Conference*, volume XXI, pages 451–454, Orlando, Florida, March 1996.

[4] D. Wiseman, J. Canaris, S. Whitaker, J. Venbrux, K. Cameron, K. Arave, L. Arave, M. N. Liu, and K. Liu. "Design and Testing of SEU/SEL Immune Memory and Logic Circuits in a Commercial CMOS Process". In *Workshop Record, 1993 IEEE Radiation Effects Data Workshop,* pages 51–55, July 1993.

[5] R. R. Troutman. *Latchup in CMOS Technology: The Problem and Its Cure.* Kluwer Academic Publishers, 1986.

[6] W. R. Dawes, Jr and G. F. Derbenwick. "Prevention of CMOS Latchup by Gold Doping". *IEEE Transactions* on *Nuclear* Science, NS-23(6):2027–2030, December 1976.

[7] J. R. Adams and R. J. Sokel. "Neutron Irradiation for Prevention of Latch-up in MOS Integrated Circuits". *IEEE* Transactions *on Nuclear Science,* NS-26(6):5069–5073, December 1979.

[8] S. C. Moss, S. D. LaLumondiere, J. R. Scarpulla, K. P. MacWilliams, W. R. Crain, and R. Koga. "Correlation of Picosecond Laser-Induced Latchup and Energetic Particle-Induced Latchup in CMOS Test Structures". *IEEE Transactions on Nuclear Science,* 42(6):1948–1956, December 1995.

[9] J. E. Schroeder, A. Ochoa, Jr., and P. V. Dressendorfer. 'Latch-Up Elimination In Bulk CMOS LSI Circuits". *IEEE Transactions on Nuclear Science,* NS-27(6):1735–1738, December 1980.

[10] J. Canaris and S. Whitaker. "Circuit Techniques for the Radiation Environment of Space". In *Proceedings of the IEEE Custom Integrated Circuit Conference,* pages 5.4. 1–5.4.4, Santa Clara, California, May 1995.

[11] D. K. Nichols, W. E. Price, M. A. Shoga, J. Duffy, W. A. Kolasinski, and R. Koga. "Discovery of Heavy-Ion-Induced Latchup in CMOS/EPI Devices". *IEEE Transactions on Nuclear Science,* NS-33(6):1696–1696, December 1986.

[12] D. Wiseman, J. Canaris, S. Whitaker, J. Gambles, K. Arave, and L. Arave. "Test Results For SEU And SEL Immune Memory Circuits". In *Proceedings of the 5th NASA* Symposium *on VLSI Design,* pages 2.6.1–2.6.10, Albuquerque, New Mexico, November 1993.

[13] S. R. Whitaker, J. W. Gambles, D. Wiseman, and G. K. Maki. "Flight Solid State Recorder Codec". In *Proceedings of the 1995 Government Microcircuit Applications Conference,* March 1995.

[14] M. White, B. Bartholet, and M. Baze. "Automated Radiation Hard ASIC Design Tool". In *Proceedings of the 5th NASA Symposium on VLSI Design,* pages 11.4. 1–1 1.4.8, Albuquerque, New Mexico, November 1993.

[15] S. E. Diehl, A. Ochoa, Jr., P. V. Dressendorfer, R. Koga, and W. A. Kolasinski. "Error Analysis And Prevention Of Cosmic Ion–Induced Soft Errors In Static CMOS Rams". *IEEE Transactions on Nuclear Science,* NS-29(6):2032–2039, December 1982.

[16] S. Whitaker, J. Canaris, and K. Liu. "SEU Hardened Memory Cells For A CCSDS Reed Solomon Encoder". *IEEE Transactions on Nuclear Science,* 38(6):1471–1477, December 1991.

[17] M. N. Liu and S. Whitaker. "Low Power SEU Immune CMOS Memory Circuits". *IEEE Transactions on Nuclear Science,* 39(6):1679–1684, December 1992.

[18] J. Canaris. "NASA SERC Digital Correlator Projects". In *Proceedings of the 4th NASA Symposium on VLSI Design,* pages 4.2.1-4.2.3, Coeur d' Alene, Idaho, October 1992.

URC97048

- 3/25

An Integrated Geoscience Analysis of Groundwater Resources in the El Paso, Juarez, Las Cruces Region

Cindy L. Gillespie, G. Randy Keller, Brian S. Penn
Pan-American Center for Earth and Environmental Studies
PACES
The University of Texas at El Paso

Introduction

Water is the most important natural resource in the El Paso/Juarez/Las Cruces region whose population is growing rapidly. These cities depend on shared aquifers for a major portion of their water. The supply of water from this resource has been rapidly declining for a number of years due to the increased demand from Texas, New Mexico, and Chihuahua. This demand for fresh water warrants further exploration and conservation activity. Information about basin fill thickness and subsurface structure is needed to understand the occurrence of ground water distribution within the known aquifers. To attack this problem, gravity data are used because they are a particularly cost-effective tool to provide information about the geometry of the basin. An overlay of the data with Thematic mapper satellite images and digital elevation models allowed us to examine the complexity of the subsurface features relative to the "simple" surface expression.

Groundwater Resources

The El Paso region is located in the southern part of the Rio Grande rift which is characterized by closed intermontane basins. The mountains surrounding the basin are the primary source of material that filled the basins during the Tertiary and Quaternary. Groundwater recharge is minor, amounting to only about 10% of present withdrawal rates. Recharge occurs through the basin sediments and also along the main mountain range fronts surrounding the aquifers in the region.

In the Hueco basin, which is El Paso's main source of groundwater, fresh water is restricted to approximately the upper 300 meters of basin fill deposits hosted mainly in the Camp Rice Formation and in the uppermost sediments of the Fort Hancock Formation. Below this zone there is a smaller lens that contains water of varying salinity (800 to 2500 TDS). It is thought that the water in the remainder of the basin is very salty with a TDS contents reaching approximately 40,000 (Client (1969).

It is estimated that for the year 2010 there will be more than 3 million people in the El Paso/Juarez/Las Cruces region. Ashworth (1990) estimated that as a consequences of this increase in population there will be a **70°/0 increase** in the water consumption in the area of which the majority will be withdrawn *from the* Hueco Basin. White (1 983) estimated that the U.S. side oft he Hueco basin had over 10 million acre feet (MAF) of recoverable fresh water remaining in storage. Forecasts of the rate of depletion (calculations based on dividing fresh water in the storage by the yearly usage) suggests drinkable ("inexpensive"-fresh) water in the Hueco Basin

will be exhausted before the year 2050. Since this is an estimate of U. S. consumption, and does not consider the rapidly rising population of Juarez, this may be an optimistic prediction.

Regional Geology

At the beginning of the Cenozoic era, compressive forces associated with the Laramide orogeny produced uplift, folding and faulting of large masses of rock in the west Texas, New Mexico, Chihuahua region. Beginning about 30 million years ago, an extensional regime associated with magmatic intrusions began affecting the area and culminated the development of the Rio Grande rift which extends from Leadville Colorado though New Mexico and El Paso into northern Chihuahua, Mexico (Ramberg et al., 1978; Keller et al., 1990). The rifting formed a series of basins (e.g., San Luis, Albuquerque, Jornada, Hueco, Tularosa, Mesilla) along the Rio Grande River. In the El Paso region, the Mesilla, Hueco, and Tularosa basins and the Franklin - Organ Mountain range represent the main structural units within the Rio Grande rift (Figure 1). The major units in the basin fill are the Fort Hancock Formation which is composed of alternating layers of clay, siltstone and minor sandstone deposited in a fluvial environment (Strain, 1966) and Camp Rice Formation (limestone conglomerate and igneous boulders cemented by caliche deposited in an alluvial environment). These older units are covered with Quaternary alluvium, terrace gravels, fan gravels, calcretes or eolian sand. The basins are expressed as broad depressions which are bounded by fairly high- angle normal faults which have downdropped the basins in relation to the adjacent mountain masses.

To date, our main emphasis has been on the Hueco basin which has a general trend north-south trend north of El Paso and northwest to southeast trend southeast of El Paso. This basin extends north into New Mexico as the Tularoso Basin. It is bounded by the Franklin Mountains on the west and by the Hueco Mountains on the east. The Tularosa basin is flanked by the Organ and Sacramento Mountains. To the south in Mexico, the Hueco basin is flanked by the Sierra de Guadalupe and the Sierra de San Ignacio.

Data Analysis and Results

Gravity data combined with remote sensing data provided an inexpensive and efficient means of evaluating the geometry of the basins. Gravity data from the UTEP database were used in this study. Standard corrections that account for known variations in gravity with latitude and elevation were applied to the data. Terrain corrections were performed, and 2.67 g/km3 was used for the crustal density. A variety of gravity maps were constructed and used as overlays to elevation models. Computer modeling of profiles of gravity data constrained by data from drill holes was employed to evaluate basin structure. We have had the advantage of a series of MS thesis completed by UTEP students on which to draw for background information (Wen, 1983: Hadi, 1991: Burgos, 1993: Imana, 1993: Lanka, 1995).

Our analysis showed that the basin structures are much more complex than their physiographic expressions would indicate. For example, instead of being associated with one continuous basin which changes trend from north south to northwest-southeast at El Paso, the Hueco basin actually consists of three sub-basins. The gravity modeling (Figure 2) shows that the main one, which lies adjacent to the Franklin Mountains, contains about 3km of Cenozoic fill. On the other hand, the Tularoso basin is divided into two sub-basins which both lie along the western

side of the physiographic expression of the basin. In this basin, the maximum thickness of-the fill is estimated to be about 2.5 km and is similar in sediment composition to the Hueco Basin.
The Mesilla basin is the southernmost basin in New Mexico. This complexly faulted basin is filled with elastic sediment, grey sandstones and shales, red to maroon silty shales, and fine grained sandstones. In the Mesilla basin region, there are a series of structural highs and lows extending from the Rio Grande River westward to Deming, New Mexico. Including pre-rift sediments and volcanics, the thickness of the basin fill is estimated from gravity data to be up to 5 km.

Summary and Conclusions

The mountains in the area are the primary source of the materials that filled the basins during the Tertiary and Quaternary. The primary role of the surrounding mountains is that they serve as a way for infiltration and recharge to the aquifers as well as a barrier to the *movement of* groundwater and a source of the dissolved solids found in it. Gravity lows exist over the basins and gravity highs are found over the mountains. Gravity lows are a result of the accumulation of less dense material in the down faulted areas. The gravity highs are the result of dense sedimentary, igneous and metamorphic rocks being brought to the surface. Small areas adjacent to the mountains are the source of fresh water aquifers which occur as small *lens* within deep aerially extensive basins. Overlays of the gravity data with the digital elevation and satellite data delineate numerous sub-basins which generally correlate with the fresh water lenses. Population growth is draining these aquifers faster than they are being replenished, therefore "inexpensive" fresh water is becoming extinct. At the present rate of extraction , the aquifers will be exhausted in about 50 years. Our results suggest the possibility of deep additional aquifers deep in the rift basin is very real and should be explored in more detail.

References

Ashworth, **J.** B., **1990,** Evaluation of groundwater resources in El Paso County, Texas, Texas Water Development Board Report 324.

Burgos, Arturo, 1993, A Gravimetric Study of the thickness of the unconsolidated materials in the Hueco Basin aquifer, Juarez area, Chihuahua, Mexico, M.S. Thesis, The University of Texas at El Paso 7-35 p.

Client, T., 1969, Groundwater occurrences in El Paso and its related geology, New Mexico Geological Society, 20th Field Conference Guidebook, p. 202-214.

Hadi, Julfi, 1991, A study of the structure and subsurface geometry of the Hueco Basin, M.S Thesis, The University of Texas at El Paso, 88 p.

Imana, Ekal, 1993, A comparative study of Lokichar, Kerio and Mesilla Basins, M.S. Thesis, University of Texas at El Paso, 25-32 p.

Keller, G. R., Morgan, P. and Seager, W. R., 1990, Crustal structure, gravity anomalies and heat

flow in the southern Rio Grande rift and their relationship to extensional tectonics, Tectonophyics, v. 174, p. 21-37.

Lanka, Kalidas, 1995, An integrated study of the Subsurface Structure of the Tularosa Basin, south central New Mexico, M. S. Thesis, The University of Texas at El Paso 4-48 p.

Ramberg I.B., Cook, F.A. and Smithson, S. B., 1978, Structure of the Rio Grande rift in southern New Mexico and west Texas based on gravity interpretation, Geological Society of America Bulletin, v.89, p. 107-123.

Strain, W. S., 1980, Pleistocene rocks in the El Paso and Hudspeth Counties, Texas, adjacent to interstate highway 10, New Mexico Geological Society, 31st Field Conference Guidebook, p. 179-182.

Wen, C. L., 1983, A study of Basin fill thickness in the southern New Mexico, west Texas and northern Chihuahua, M. S. Thesis, University of Texas at El Paso, 74 p.

White, D. E., 1983, Summary of Hydrologic information in the El Paso, Texas area, USGS open file report, 83-775, 77 p.

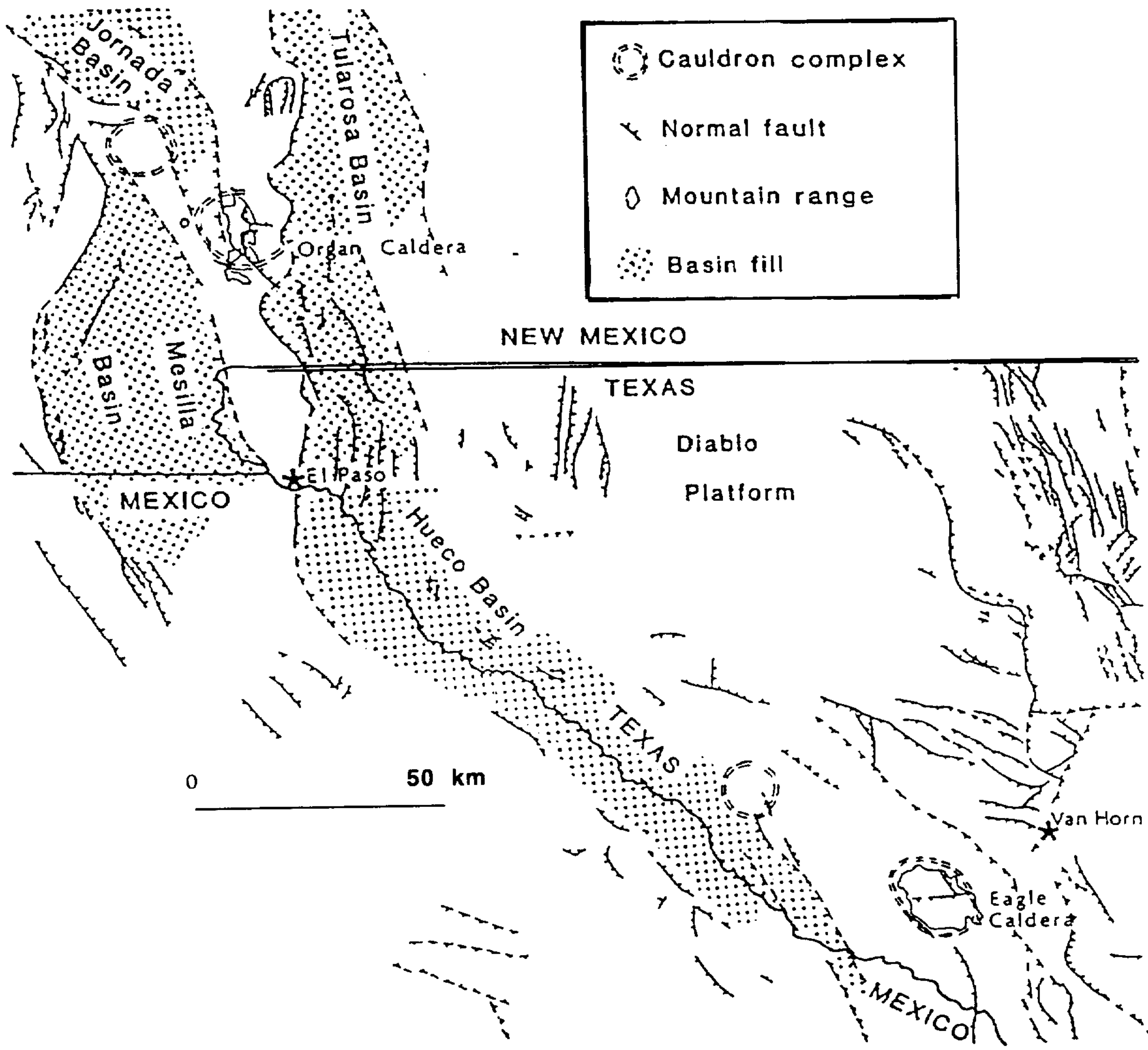


Figure 1
Index map of the El Paso, Juarez, Las Cruces region. Mountain ranges that flank the Hueco basin are the Organ, Franklin and Sierra de Juarez, to the west and the Hueco Mountains to the east. The Hueco Basin is divided into three sub basins (Tularosa, Hueco, and Mesilla). The Hueco basin extends north into New Mexico as the Tularosa basin and south into Mexico. In the south, *the basin* is bounded by the Sierras de Guadalupe and San Ignacio, and the Quitman, Malone and Finlay Mountains.

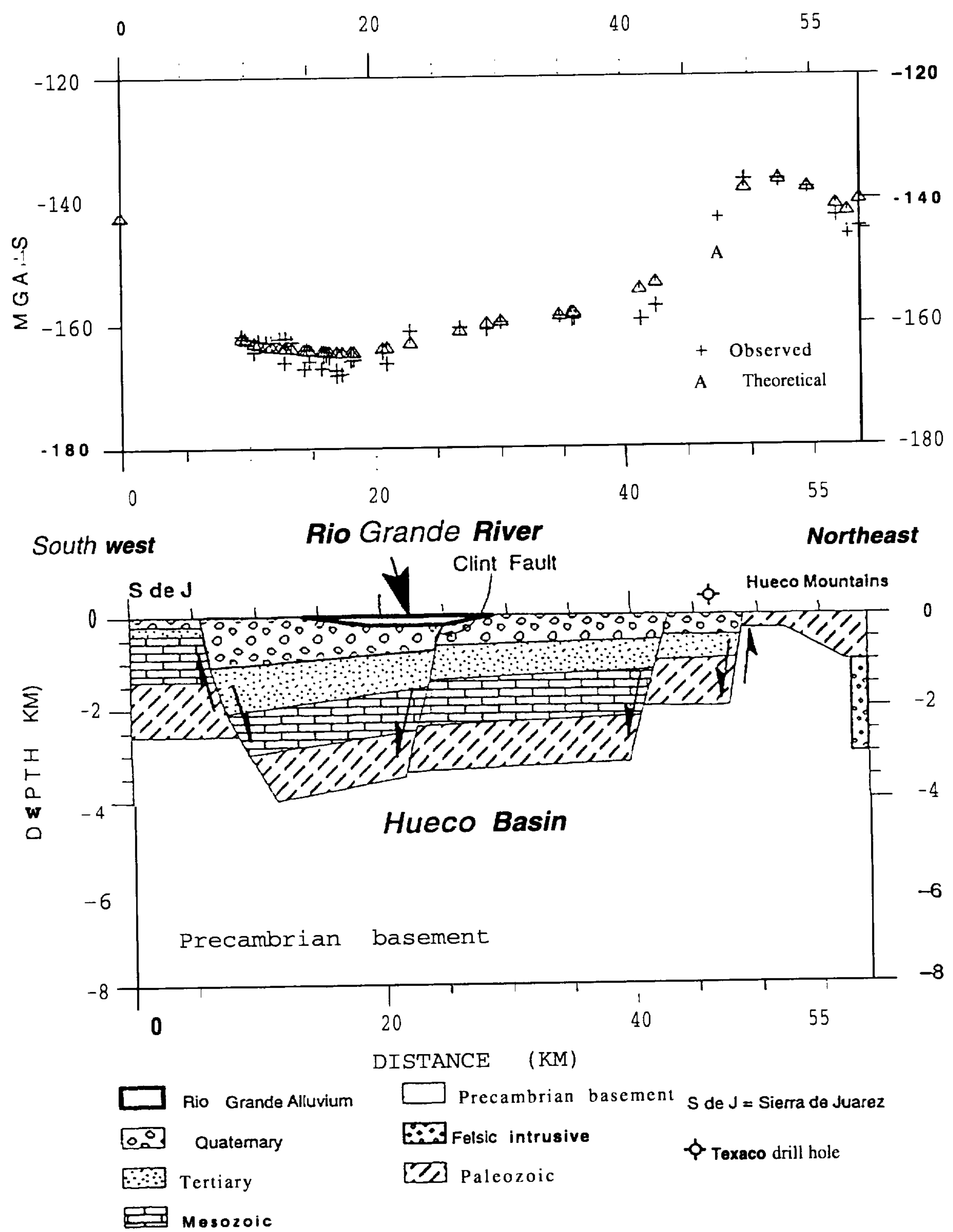


Figure 2
Computer model of a gravity profile crossing the Hueco basin from near Juarez on the southeast to the Hueco Mountains west of El Paso.

URC97049

N A S A
54-7/43

ATMOSPHERIC CORRECTION OF SATELLITE IMAGERY USING MODTRAN 3.5 CODE*

Fabián O. González and Miguel Vélez-Reyes
Advanced Automated Image Processing Project
Tropical Center for Earth and Space Studies
University of Puerto Rico-May agüez
Mayagüez, Puerto Rico 00681
Tel, (787) 832-4040, x3086, 3094 Fax. (787) 831-7564
e-mail: fabiang@exodo.upr. clu.edu mvelez@exodo.upr. clu.edu

Abstract—When performing satellite remote sensing of the earth in the solar spectrum, atmospheric scattering and absorption effects provide the sensors corrupted information about the target's radiance characteristics. We are faced with the problem of reconstructing the signal that was reflected from the target, from the data sensed by the remote sensing instrument. This article presents a method for simulating radiance characteristic curves of satellite images using a MODTRAN 3.5 band model (BM) code to solve the radiative transfer equation (RTE), and proposes a method for the implementation of an adaptive system for automated atmospheric corrections. The simulation procedure is carried out as follows: (1) for each satellite digital image a radiance characteristic curve is obtained by performing a digital number (DN) to radiance conversion, (2) using MODTRAN 3.5 a simulation of the images characteristic curves is generated, (3) the output of the code is processed to generate radiance characteristic curves for the simulated cases. The simulation algorithm was used to simulate Landsat Thematic Mapper (TM) images for two types of locations: the ocean surface, and a forest surface. The simulation procedure was validated by computing the error between the empirical and simulated radiance curves. While results in the visible region of the spectrum where not very accurate, those for the infrared region of the spectrum were encouraging. This information can be used for correction of the atmospheric effects. For the simulation over ocean, the lowest error produced in this region was of the order of 10^{5} and up to 14 times smaller than errors in the visible region. For the same spectral region on the forest case, the lowest error produced was of the order of 10-4, and up to 41 times smaller than errors in the visible region,

1. INTRODUCTION

In the solar spectrum, sensors on Earth remote sensing satellites measure the radiance reflected by the atmosphere-Earth surface system illuminated by the sun. While this signal depends on the surface reflectance, it is perturbed significantly by two atmospheric processes, the gaseous absorption and the scattering by molecules and aerosols[1]. Atmospheric molecules and aerosols can modulate the radiation reflected from the earth by attenuating it, changing its spatial distribution, and introducing into the field of view, radiation from sunlight scattered in the atmosphere This provides the sensor, corrupted information about the target's radiance characteristics, resulting in blurred images or altered contrasts in image colors. We are faced with the problem of reconstructing the signal that was reflected from the target from the data sensed by the remote sensing instrument. This is not generally a simple process due to the spatial and temporal dependence of the atmospheric effects.

Simulation of satellite imagery is a powerful tool that can be used effectively in the development and implementation of preprocessing techniques on the digital data. Results from simulations of sensor measured radiance can be used in the development of atmospheric correction algorithms to be applied to satellite data, Our goal in this research is the development of an adaptive system for automated atmospheric corrections of the digital data. The information obtained from the simulation (i.e., from the solution of the RTE) provides an estimation of the corrected radiance. This problem can be treated as a filtering system (see Figure 1). The spatial and temporal variability of atmospheric effects suggests the study of adaptive systems and approaches to on-line adaptation of atmospheric models used in atmospheric corrections. The estimated surface reflectance extracted from the radiation received at the sensor, can be processed by some adaptation mechanism to estimate the atmospheric parameters and readapt the model for new computations of the RTE. The automated and adaptive nature of this system should make it a powerful one.

* Sponsored by NASA URC Program under grant NCCW-0088.

We selected a Landsat TM satellite image with no atmospheric corrections, to simulate its spectral radiance characteristic. TM images contain digital data for seven channels or spectral bands located in different atmospheric windows (see Table 1). This gives us information about the radiance in the visible region of the spectrum, and the near, mid, and thermal infrared. We selected two surfaces with known reflectance characteristics to simulate: a forest, and the ocean. The simulations were carried out using MODTRAN 3.5, This is a computer code based in a subroutine of the equation of radiative transfer, capable of making calculations of atmospheric transmittance or radiance at a mid spectral resolutions. The code was developed by researchers from Phillips Laboratory, at Hanscom Air Force Base in Boston. Massachusetts.

2. PROCEDURE

Using a computer software for image processing, we obtained the digital number (DN) for a single pixel in each of the surfaces, and for each channel. We collected seven DN values for each surface, i.e., one value for each channel. Conversion from the DN of TM digital data to at-satellite spectral radiance was done using calibration data supplied by EOSAT. After this conversion was performed, we had seven radiance values for each scene simulated. One value for each one of the TM channels, Plotting these values against the central wavelength of the sensors' bandwidths, we obtained measured radiance curves for surfaces of forest and ocean.

We executed MODTRAN 3,5 setting the input parameters to simulate the surface, atmospheric, and geometry characteristics of the real system. By executing the program we obtained top-of-the-atmosphere (TOA) spectral radiance, or at-satellite spectral radiance, for a spectral band from 0.4μm to 12.5 μm. We compared the curves obtained using MODTRAN 3.5 with those obtained from the TM digital data.

Table 1 Atmospheric Windows for Terrestrial Observations by Satellites

Spectral Region	Atmospheric Window (pm)	TM Sensor Band (μm)
Visible	(between) 0.4-0.75	band 1:0.45-0,52
		band 3:0.63-0.69
Near Infrared	(at about) 0.85	band 4:0.76-0.90
Middle Infrared	(at about) 1.06, 1.22, 1.60	band 5:1.55-1.75
	(at about) 2.20	band 7:2.08-2.35
Thermal Infrared	(at about) 4.00, 11.00	band 6:10.40-12.50

We can summarize this procedure in the following steps:

1. Select an appropriate satellite image in digital data. We will use pixels from this image for the simulation.
2. Select a number of pixels in the image with known reflectivity and emissivity characteristics, which correspond to different type of surfaces, one pixel per type of surface to be used in the simulation.
3. Obtain the digital numbers (DN) for each selected pixel.
4. Repeat step 3 for each channel available in the digital data.
5. Convert each of the DN values to spectral radiance values.
6. Run the MODTRAN 3.5 simulation, setting the input parameters so that they match as close as possible, the digital data used for the simulation.
7. Compute the average radiance for the spectral range that corresponds to each of the satellite sensors,
8. Plot these average values against wavelength, using the center wavelength of each spectral band,
9. Repeat steps 8 through 10 for each case emulated.
10. Plot the converted DN to radiance values from the digital data, using the same wavelength axis as in MODTRAN plots.
11. Calculate the error between both sets of graphs.

3. EVALUATION OF RESULTS

3.1. *DN to Radiance Conversion of the TM Digital Data*

The curves corresponding to the DN to radiance conversion of the ocean and forest surfaces are plotted in Figure 2a and 2b, respectively. Knowing that a change in DN implies a change in radiance, we can see how this radiometric quantity varies with wavelength in each selected surface.

Looking at the curve for the ocean surface, we would expect water to show high absorption in the near and mid-IR spectral region[4]. This is probably the most distinctive reflectance characteristic of water. We can see from Figure 2a, that at these wavelengths the lowest radiance values (and generally lower DN values) are recorded. Atmospheric scattering and absorption effects are mostly present in the visible region of the region of the spectrums. The curve for the forest surface differs greatly from the curve for ocean in the visible region of the spectrum. In the next section we will see how the simulation helps us in the analysis and interpretation of the DN to radiance conversion curves, especially in this spectral region.

3.2. *MODTRAN 3.5* Simulation

MODTRAN simulates the most important radiance characteristics for a given atmosphere-Earth system. These are: atmospheric transmittance, optical depth, path thermal radiance, surface emission, total path scattered radiance, ground reflected radiance, and total radiance[6]. We will limit our discussion to the total radiance and the ground reflected plus surface emitted radiance. This will suffice to show the dramatic effects of atmospheric absorption and scattering in solar radiance, and to show how the simulations provide a great deal of information to be used when developing correcting tools.

The curves of the ground reflected plus surface emitted radiance are plotted in Figure 3a and 3b, for ocean and forest respectively. These curves approximate what we would receive at the satellite sensor if scattering effects where not present. The ground reflected radiance is constituted by a direct reflection term and a scattered reflected term. The direct ground reflected includes the directly transmitted solar radiance reaching the ground, whereas the scattered term is the solar radiation scattered within the atmosphere before reaching the ground. This scattering effect is increased in the visible region of the spectrum.

Now knowing what is the desired response, let us take a look at the curves of the total radiance in Figures 4a and 4b. We would like to eliminate the scattered ground reflected radiance, path scattered radiance, and path emitted radiance. The fact is that these terms are included in the total radiance. Thus, in the figure, instead of having a measurement of true surface reflected and emitted radiance, we have obtained a false or apparent radiance'. This leads to an altered or apparent reflectance characteristic of the sensed surface. We have to mention that the molecular absorption effects have been diminished by measuring the radiance in adequate atmospheric windows. Still, we are not dealing in any way with the measurement of aerosol absorption effects,

The difference between the total radiance measurement and the ground reflected and emitted radiance is shown in Figures 5a and 5b.

3.3. *Mean Radiance Values for Spectral Bands Corresponding to TM Channels*

The purpose of' calculating the mean radiance for seven spectral bands is to approximate the radiance measurements as they would be obtained with the Landsat TM sensors. To do this, we are assuming that the radiance measured by one of the TM sensors is the average value of the point of interest plus the values of its neighbors. Thus, by calculating the mean radiance value over the selected spectral bands, we will be simulating the total radiance measured by the sensors. We also assume that this measured radiance corresponds to the center wavelength of each spectral band.

Figures 6a and 6b contain the mean values for the ground reflected plus surface emitted radiance. Figures 7a and 7b show the mean total radiance per channel calculated from the simulation. Figures 6 and 7 derive from Figures 3 and 4, respective] y. The difference **between** Figures 6 and 7 is the undesired radiance introduced by the atmospheric effects.

3.4. *Comparing the empirical and simulation results*

We compare the results based on the error between plots for DN to radiance conversion and mean values calculations for total radiance, i.e., the difference between Figures 2 and 7. The curve in Figure 8a is the error produced *in* the simulation of the ocean surface. From this figure we see that the minimum error occurs for channel 7 in the mid-infrared (2.08 -2.35 pm), and it is of the order of 10-5. The maximum error occurs in channel 3 (red, 0.63- 0.69 μm) in the visible region. This error is of the order of 10^{-4} . Exacty, the error in channel 3 is 14 times higher than the error in channel 7. In general, the error in the visible region of the spectrum will be the highest. We can mention a set of factors that could influence the outcome of this simulation.

We know that scattering effects are more critical within the visible region of the spectrum. The measured radiance will depend, to a high extent, on the atmospheric characteristics of the radiance path. In other words, there will be a significant component of the total radiation that will always vary according to the location where a certain surface is sensed. This alone introduces a high degree of uncertainty, The simulation errors can be minimized by using atmospheric and surface models which are representative of the surface to be simulated, and of the atmosphere

over that surface. We cannot have general models that work all the times, especially within the visible region of the spectrum. This addresses the issue of an adaptive correction system.

Since the total radiance of the simulation approximates the empirical results, we can estimate the correct radiance in the image by subtracting the difference between Figures 6 and 7 from Figure 2. Conceptually, we can use the surface and atmospheric parameters generated by the simulation to implement an adaptation mechanism for the correction of the measured radiance (Figure 1).

Figure 8b presents the error in the simulation of the forest surface, Maximum and minimum errors occur in the visible and thermal-IR respectively. We would expect the radiance in the blue and red portions of the spectrum to be lower than radiance in the green channel. Looking at Figure 8b we will notice the peak in the blue channel corresponding to TM channel 1. The reason for this might reside on the selection of reflectivity and emissivity for forest. The available forest reflectivity and emissivity values may not be representative of the type of forest sensed in the image. The chlorophyll content of the plant leaves will affect the reflectivity in the visible region of the spectrum. Thus, reflectivity in the visible may vary greatly as the type of vegetation that grows in a particular forest changes[4].

For wavelengths in the infrared, the error between the radiance magnitudes for a particular wavelength is much smaller than the error for wavelengths in the visible. We see from the errors in channels one and seven that the error corresponding to the blue spectral region (0.45 -0.52 μm) is 41 times higher than the error in the thermal-IR, where the error is of the the order of 10-4. The small errors in the IR region indicate the existing similarity between the image and the simulation, not only in the proportionality within each case, but also in the magnitude of the values.

4. SUMMARY AND CONCLUSIONS

We have presented a simple but useful method for the simulation of solar radiance, and its potential applications to the correction of atmospheric effects in satellite imagery. By simulating effects of scattering and absorption in solar radiance, atmospheric and surface parameters can be obtained that can be used to estimate correct surface radiance. Simulations based on the radiative transfer equation (RTE) can be used to estimate radiance values to be compared to values measured by satellite sensors. The error between these values can be used in some adaptation mechanism, which will generate new atmospheric and surface parameters to adapt the model for new radiative transfer computations.

We used MODTRAN 3.5, a code based on a subroutine of the RTE to simulate radiance characteristics of ocean and forest surfaces in Landsat Thematic 'Mapper (TM) digital data. In general, higher errors between the simulated and measured values were obtained in the visible region of the spectrum, due to greater atmospheric scattering in this spectral region. Within the spectral range from near-IR to far-IR, results of the simulation were very encouraging. For the ocean surface, simulation errors in the visible region of the spectrum were up to 14 times higher than errors in the IR spectrum, were the minimum errors were of the order of 10^{-5}. For the forest surface, errors were of the order of 10"4 on the IR region, and errors in the visible region were op to 41 times higher.

The simulation process still needs to be fine tuned. Uncertainty factors must be minimized to assure the accuracy of the simulations, especially in the visible region of the spectrum. This can be clone by adding to the simulation code, atmospheric and surface data that models more accurately the real case. Also, we need to address computational issues. The procedure developed is applied in a pixel by pixel basis. Application of this procedure on blocks of pixels in an image, must take into consideration the implementation of some architecture to perform parallel processing of the pixels.

5, REFERENCES

1. Vermote E., Tanré D., Deuzé J. L., Herman M., and Morcrette J.J. **Second Simulation of the Satellite Signal in the Solar Spectrum (6 S).** Laboratoire d'Optique Atmosphérique, Université des Sciences et Technologies de Line, Line, France, 216 pp.
2. Asrar, G. 1993. Theory **and Applications of Optical** Remote Sensing. John Wiley and Sons, New York, USA, 730 pp.
3. Berk, A., L.S. Bernstein, D.C. Robertson, 1989, **MODTRAN: A** Moderate Resolution **Model** for **LOWTRAN** 7. Phillips Laboratory, Hanscom US Air Force Base, 38 pp.
4. Lillesand, T.M., and R.W. Kiefer. Remote Sensing and Image Interpretation. John Wiley& Sons, NY, USA.
5. Lenoble, J. Atmospheric Radiative Transfer. A. DEEPAK Publishing, Hampton, Virginia USA, 532 pp.
6. Berk, A., L.S. Bernstein, D.C. Robertson, 1995. **MODTRAN** 3 User Instructions. Phillips Laboratories, Hanscom US Air Force Base, 38 pp.

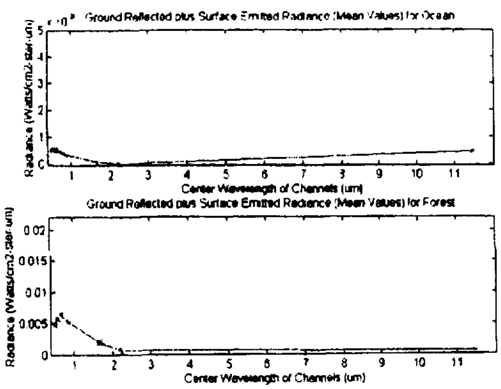


Figure 6 Channel Mean Values for Ground Reflected plus emitted radiance

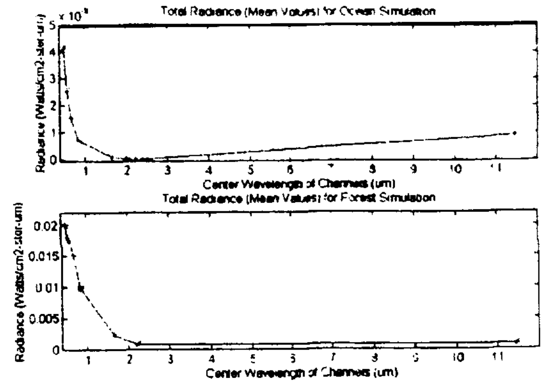


Figure 7 Channel Mean Values for Total Radiance

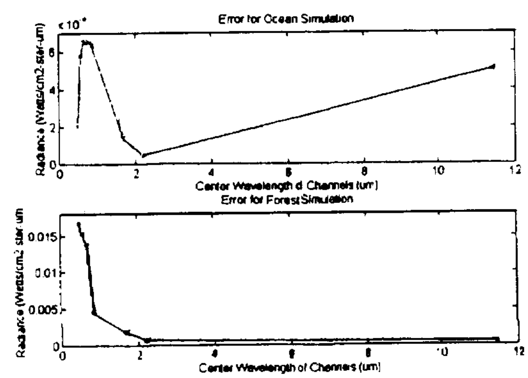


Figure 8 Final error for simulated cases

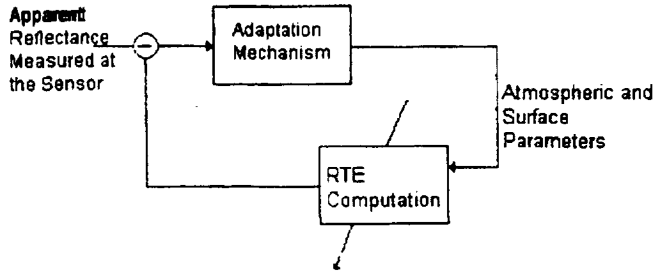


Figure 1 Model of an adaptive system for automated atmospheric corrections

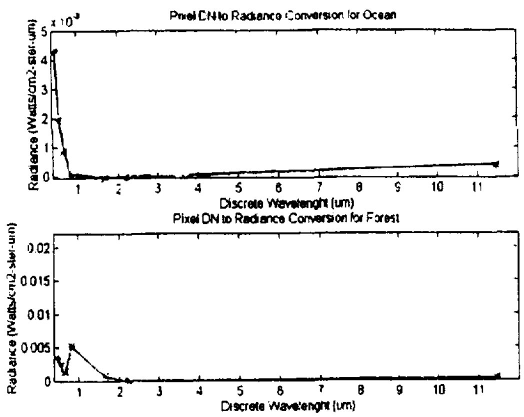


Figure 2 DN to Radiance Conversions

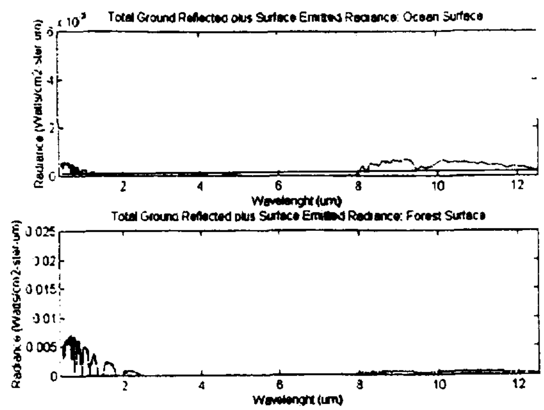


Figure 3 Radiance due to Ground Reflection and Emission

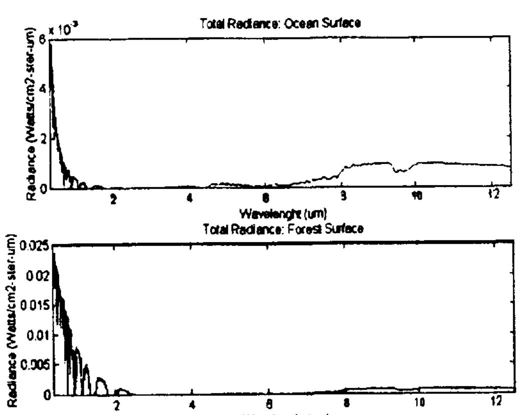


Figure 4 Total Radiance

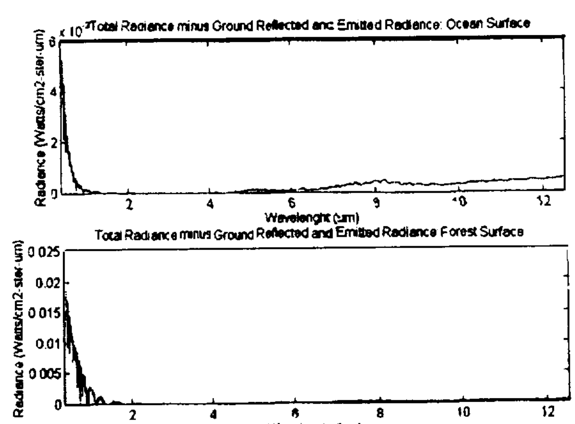


Figure 5 Radiance due to Scattering Effects

Gas Dynamics, Characterization, and Calibration of Fast Flow Flight Cascade Impactor Quartz Crystal Microbalances (QCM) for Aerosol Measurements.

J. R. Grant, A. N.Thorpe, C. James, A. Michael, M. Ware, F. Senftle, and S. Smith
Center for the Study of Terrestrial and Extraterrestrial Atmospheres
Howard University, Washington, D.C. 20059

During recent high altitude flights, we have tested the aerosol section of the fast flow flight cascade impactor quartz crystal microbalance (QCM) on loan to Howard University from NASA. The aerosol mass collected during these flights was disappointingly small. Increasing the flow through the QCM did not correct the problem. It was clear that the instrument was not being operated under proper conditions for aerosol collect ion primarily because the gas dynamics is not well understood. A laboratory study was therefore undertaken using two different fast flow QCM's in an attempt to establish the gas flow characteristics of the aerosol sections and its effect on particle collection, Some tests were made at low temperatures but most of the work reported here was carried out at room temperature.

The QCM is a cascade type impactor originally designed by May (1945) and later modified by Anderson (1966) and Mercer et al (1970) for chemical gas analysis. The QCM has been used extensively for collecting and sizing stratospheric aerosol particles (Chuan and Woods, 1984; Woods and Chuan, 1983; and Chuan et al, 1981). In this paper all flow rates are given or corrected and referred to in terms of air at STP. All of the flow meters were kept at STP. Although there have been several calibration and evaluation studies of moderate flow cascade impactors of less than or equal to 1 L/rein. (Marple, Liu and Whitby, 1974), there is little experimental information on the gas flow characteristics for fast flow rates greater than 1 L/rein.

The Quartz Crystal Microbalance

To ensure the results were not merely artifacts of a particular instrument, these studies were carried out with the NASA 6-stage aerosol section QCM and also a 10-stage aerosol QCM designed by California Measurements Corp (CMC).

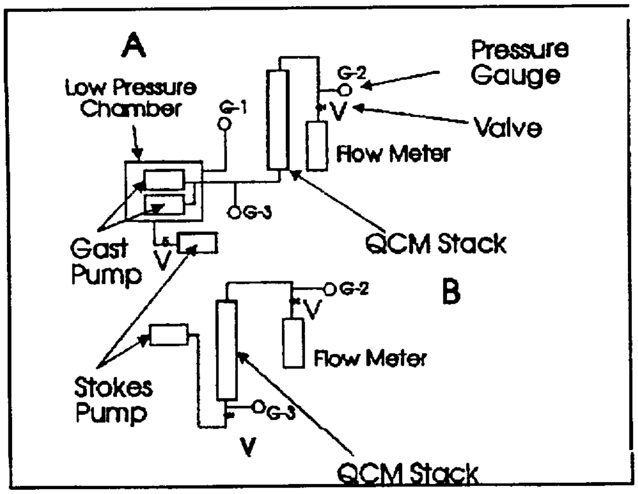
The NASA **QCM:** Each stage of this instrument is comprised of an inlet nozzle and a cylindrical housing for two quartz crystal oscillators and their associated electronics. The inlet nozzle is actually mounted in the bottom of the previous stage. When a stage is removed from the stack both its nozzle in the previous stage and housing are removed. The inlet gas and particles impinge on the upper quartz crystal, and any particulate matter in the gas may or may not adhere to the quartz surface and change its vibration frequency. The lower crystal oscillator is used for reference to produce a change in beat frequency and to compensate for frequency shifts due to temperature changes. As the upper crystal picks up particles, the beat frequency between the two crystals changes with the increase in mass.

The California Instruments **QCM:** This instrument is generally similar to the NASA QCM but has 10 stages and was calibrated by Hering (1987) for stratospheric sampling using flow rates of 1.2 L/rein. Each stage is 3 inches in diameter compared to a diameter of 1 3/4 inches for the NASA QCM. Unlike the NASA instrument, the inlet jet for each stage is located in the stage housing.

Experimental Method

The first set of experiments were performed with the NASA QCM. in order to simulate flight conditions the QCM was set up as shown in Figure 1A. Two GAST vacuum pumps were mounted in a large partially evacuated chamber which was used to simulate the low pressure condition in the stratosphere. Varian electronic pressure gauges (Model # WV100-2, designated G-1, G-2, and G-3) and a MKS flowmeter (Type 0558A-050L-SV)were arranged as shown in the sketch. To simulate operation in the stratosphere, the inlet pressure and the chamber pressure were kept the same. The pumps evacuated the QCM directly, and it was observed that the exit pressure of the QCM was always about 28% of the inlet pressure. At any particular pressure the flow rate did not change when the pressure difference across the QCM was further increased. Therefore, the system was always in the choke flow mode of operation. It is clear that under flight conditions the instrument did not operate properly, and why it failed to collect additional aerosol particles with increase of flow rate.

To study the gas dynamics, a much simpler arrangement was used (Figure 1 B). In order to control the exit pressure the chamber was eliminated. The inlet and exit pressure could be controlled by the two valves, the pressure difference across the QCM could be varied, and the gas flow could be controlled from about zero to choke flow.



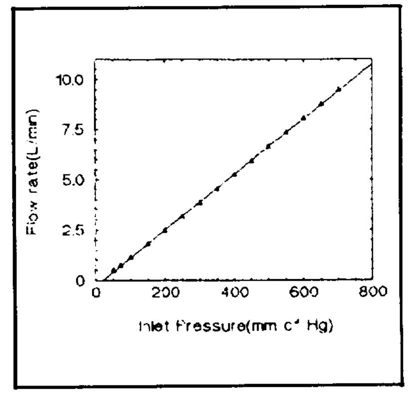
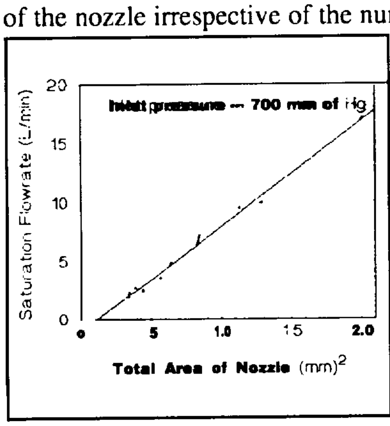


Figure 1- A- Sketch of experimental configuration used to simulate flight operation of the QCM. B- Laboratory experimental configuration to study the gas dynamics of the QCM.

Figure 2- The gas flow rate as a function of the inlet pressure for the NASA 6-stage QCM under simulated flight conditions.

Figure 2 shows how the gas flow through the 6-stage NASA QCM varies with inlet pressure using the experimental configurate ion shown in Figure 1A. As this configuration simulates the flight conditions, it is obvious that during flight the flow rate was linearly related to the pressure difference across the QCM. As the QCM is operating in the choke flow mode, such a linear relationship is expected. The choke flow is due to the nozzle with the smallest jets i.e. with the smallest total area. Figure 3 shows how the flow rate changes with the total area of the individual nozzles in the QCM's. As expected the saturation or choke flow is linearly related to the total area of the nozzle irrespective of the number of the jets per nozzle of the QCM.



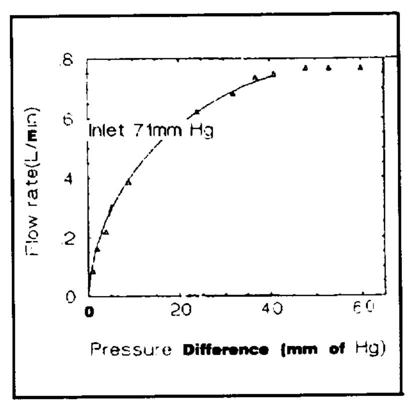


Figure 3- Saturation flow rate as a function of the total area of the individual nozzles.

Figure 4- Gas flow rate as a function of the change in 'pressure across the NASA-QCM for an inlet pressure of 71 mm of Hg.

To obviate choke flow, a series of tests were made using the configuration in Figure 1 B, the configuration which should be used in further flights for aerosol collection. At fixed inlet pressures of 71, 100, 300, 500, 650, and 700 mm of Hg the flow rate was measured as a function of the pressure difference across the 6-stage QCM. Figures 4 and 5 are typical of the flow as a function of the pressure difference across the QCM for inlet pressures of 71 and 100 mm of Hg, respectively. The data shows saturation or choke flow for large pressure differences. A similar study was made on the CMC-QCM (Fig. 6). Choke flow occurs at a lower flow as would be expected because the area of each nozzle is much less than the nozzle areas for the NASA QCM.

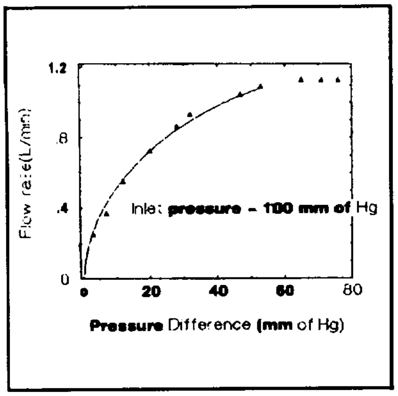


Figure 5- Gas flow rate as a function of change in pressure across the NASA-QCM for an inlet pressure of 100 mm of Hg.

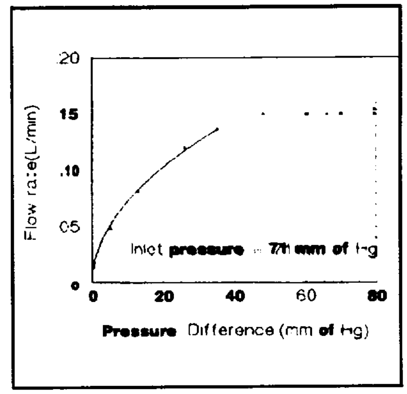


Figure 6- Gas flow rate as a function of change in pressure across the CMC-QCM for an inlet pressure of 71 mm of Hg.

The critical dimensions and other characteristics of each stage in the NASA QCM are given in Table 1. The jet-to-quartz crystal distance, S, is approximately one half of the jet diameter, W, for each stage, whereas the ratio of the throat-length, T, to the jet diameter is different for each stage. Similar data for the CMC 10-stage QCM are given in Table 2. The critical parameters for the QCM round jet impactors shown in the Tables 1, and 2, are calculated using the following relations.

(1) Effective cut-off aerodynamic diameter, ECAD (Mercer and Stafford 1969) for the aerosol particles, is

$$\mathrm{ECAD}\,(\mu m) = 1.257 \times 10^{3}[W^{3}/F]^{1/2},$$

where W is the jet diameter in cm and F is the volumetric flow rate in cm^3/min. per jet.

(2) The flow Reynolds number, N_{RE}, for a circular jet is given by

$$N_{RE} = \rho W v/\eta$$

where p is the physical density of the air, η is the viscosity of air, and v is the air flow velocity expressed in terms of the volumetric flow rate and j et diameter:

$$v = 4F/\pi n W,$$

where n is the number of jets per nozzle.

(3) The collection efficiency of the impaction stages were determined with respect to the dimensionless Stokes number (See Fairchild and Wheat, 1984), defined as:

$$S_{tk} = \rho v(\mathrm{ECAD})^{2}/9\eta w$$

where p = 1 for unit density particles. The calculations were computed using the constants

p = 1.293×10^{-3} g/cm^3 physical density of air (particles)

η = 1.82×10^{-4} viscosity of air at 20 “C

Table 1

Characteristic parameters of the NASA 6-stage QCM at a volume flow rate of 7.5 L/rein (at standard temperature and pressure)

Stage Number	No. of Jets	Jet Diameter W (mm)	Throat Length T (mm)	Jet to Crystal Separation S (mm)	S/W	T/W	ECAD (pm)	N_{Re}	Stokes #
1	1	6.6	7.11	3.31	0.502	1.077	7.783	1715	0.2038
2	1	4.2	6.83	2.17	0.517	1.626	3.95	2692	0.2034
3	1	2.88	9.06	1.39	0.483	3.146	2.25	3813	0.2045
4	2	1.72	3.36	0.97	0.564	1.953	1.464	3286	0.2023
5	4	0.79	2.33	0.42	0.532	2.949	0.644	3577	0.203 I
6	4	0.60	1.89	0.30	0.500	3.15	0.426	4711	0.2029

Table 2
Characteristic parameters of the California Measurements Corp. 10-stage Q M at a v lume flow rate of 1.5 L/rein (at standard temperature and pressure)

Stage Number	No. of Jets	Jet Diameter W (mm)	Throat Length T (mm)	Jet to Crystal Separation S (mm)	S/W	T/W	ECAD (μm)	N_{Re}	Stokes #
1	1	6	11.20	3.35	0.56	1.87	15.086	377	0.2035
2	1	4	12.30	2.25	0.56	3.08	8.2118	565	0.2034
3	1	3	13.00	1.50	0.50	4.33	5.334	746	0.2012
4	1	2	13.00	1.55	0.77	6.50	2.903	1130	0.2033
5	1	1.392	12.95	1.55	1.11	9.30	1.682	1624	0.2024
6	1	1.005	14.00	0.50	2.01	13.93	1.034	2250	0.2033
7	1	0.689	14.10	0.45	0.65	20.49	0.586	3280	0.2028
8	2	0.494	14.25	0.30	0.61	28.87	0.5049	2287	0.2031
9	4	0.331	14.20	0.20	0.60	42.87	0.391	1706	0.2031
10	7	0.249	14.40	0.15	0.60	57.88	0.335	1303	0.2029

Study of Individual Stages

As the number of jets and jet diameters vary between different QCM stages, it is desirable to know how the flow varies with pressure difference for individual stages. This was accomplished by removing each stage consecutively starting with stage # 6 and its inlet nozzle; then stage #5, and so on until all stages were removed. The pressure drop across individual stages is calculated as follows. For example with the stage #6 and its inlet jet removed from the stack, the exit pressure of stage #5 is determined. This is the same as the inlet pressure to stage #6, if stage #6 were in the stack. The exit pressure for the entire stack has been previously determined, and therefore the pressure drop across stage #6 when it is in the complete stack is the difference between the exit pressure for the 5-stage stack and the exit pressure for the 6-stage stack. The pressure difference across each stage in the stack was determined by using the same method. The sum of the pressure drop across the individual stages was equal to the pressure drop across the complete stack. As the flow rate is established by the pressure difference across the complete stack for a given inlet pressure, it is to be noted that the inlet pressure to a stage is a function of the flow rate. This is true because there is a pressure drop proportional to the flow rate across the preceding stages. Fig 7a shows the inlet pressure to stage six (outlet pressure for stage 5) as a function of the flow rate for a stack inlet pressure 71 mm Hg with stage #6 removed. The pressure difference versus flow rate for stage #6 at 71 mm Hg is shown in Fig. 7b.

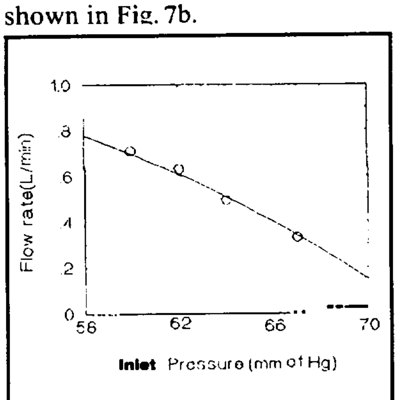


Fig. 7a Flow rate versus Inlet pressure to stage # 6 at 71 mm Hg

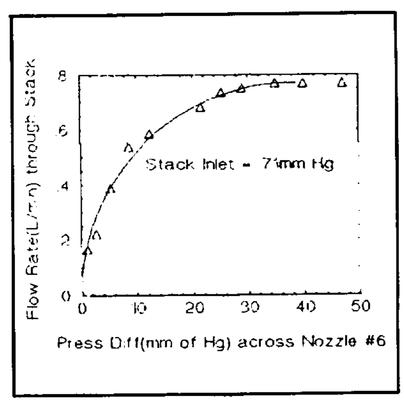


Fig. 7b Flow rate versus pressure difference for stage #6 at 71 mm Hg.

Smoke Experiments

Using both QCM's several smoke sampling experiments were conducted using the complete stacks. Smoke was introduced simultaneously to the CMC and the NASA six-stage QCM's at atmospheric pressure. No filter was used and the air flow was set at approximately 75 % of choke flow rate; that is 7.5 L/rein for the NASA stack and 1.5 L/rein for the CMC stack. These are the same flow rates used in the tables. Typical beat frequency response curves are shown for stage 5 (NASA-QCM) and stage 7 (CMC-QCM) in Fig. 8.

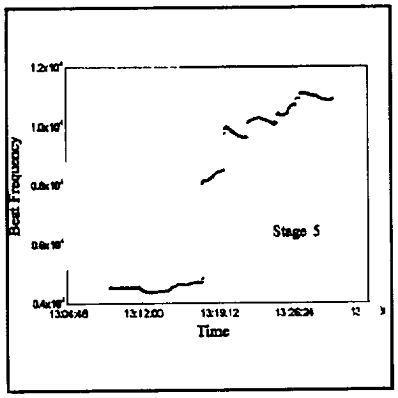


Figure 8a Beat frequency response of the crystals on stage 5 versus collection time for the NASA QCM.

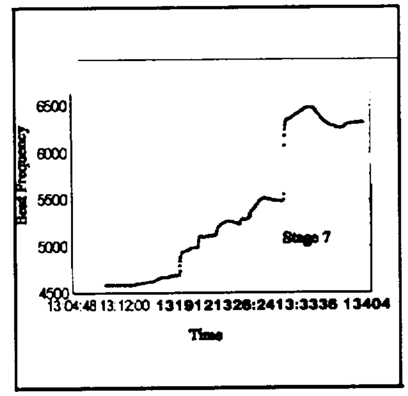
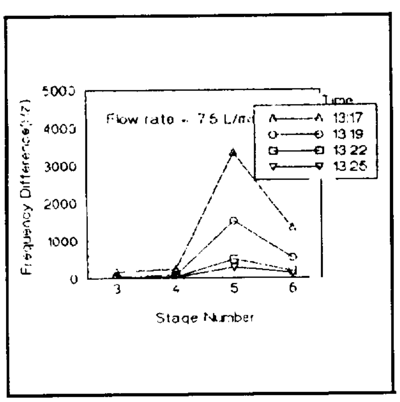


Figure 8b - Beat frequency response of the crystals for stage 7 versus collection time for the CMC-QCM.

Note the sharp increase in beat frequency at the specific time at which the smoke was introduced. The sharp increase in beat frequency is due to the accumulation of smoke particles on the quartz crystals. The total mass of particles collected is reflected in the frequency difference before and after the introduction of the smoke. The frequency difference is shown as a function of the stage number in the respective stack in Figure 9. Stage #5 and stage #7 in the NASA and CMC stacks respectively accumulated the largest mass of aerosol particles. In Tables 1 and 2, these stages correspond to an effective cut-off aerodynamic diameter of 0.644 and 0.586 μm respectively. Thus, within experimental error those stages with the same ECAD in their respective QCM's are collecting approximately the same size particles. To a first approximation the relative sensitivity is related to the flow: i.e. the frequency change for the NASA QCM is approximately 5 times that of the CMC-QCM



'igure 9a

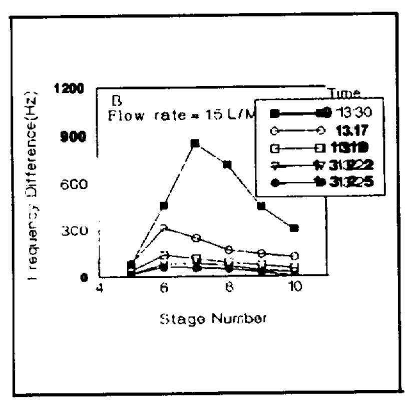


Figure 9b

Figure 9- Cumulative mass collection measured as a frequency difference for the complete stack versus the number of stages for the NASA(a) and CMC(b)QCM's.

In both figures 8a and 8b after the increase due to the smoke introduction, there is almost always a steady decrease in frequency for both QCM's. Further study of this phenomena showed that if the flow rate was kept constant, using

clean air after smoke introduction, the frequency would decrease almost to the original value before the introduction of smoke. This decrease occurs because the particles are being blown off the crystal. In the previous flight experiments the crystals were not grease coated and the pumps were left on after sampling with filtered air being pumped through the QCM until the aircraft landed and therefore much of the collected aerosol mass probably blew off the crystals. in recent laboratory experiments the crystals were coated with grease, but the grease also blew off in both QCM's at high mass flows, However, it appears that very little grease blows off at a QCM inlet pressure of 100 mm of Hg or less at these high flow rates.

Discussion

The QCM is an instrument that is used extensively in aerosol research. However, to use it at high mass flow rates one has to alter standard procedures. Recently Chuan (1993) has successfully used a four stage QCM for stratospheric aerosol measurements at a flow rate of 1.3 L/min. Our experiments indicate that a higher flow rate is possible if one designs the nozzles carefully (i.e. using the proper number of jets to maintain a low mass flow.) and keeps the flow within the correct limits at low pressures, or use an appropriate amount of grease coating so that part of the accumulated mass does not blow off the quartz crystals. Experiments are continuing on the low pressure flow rates and amount of grease coating to be used on the crystals for different experiments.

Acknowledgements

We gratefully acknowledge the financial support of Howard University and the Center for the Study of Terrestrial and Extraterrestrial Atmospheres funded by NASA grant NAGW-2950. We would like to thank R. Chuan, D. Woods, and W, Chiang for their support and many helpful discussions and I. Heard for helping with the data analysis.

References

Anderson, A. (1966) Am. Ind.Hyg.Ass.J. 27, 160
Chuan, R. L., Woods D.C. and Mc Cormick, M.P. (1981) Science 211: 830-832
Chuan, R. L., (1993) Atmos. Environ. Vol 22A. No. 17/18 pp 2901-2906
Chuan R.L. and Woods D.C. (1984) Geophys. Res. Lett 11:553-556
Fairchild, C.I. and L.D. Wheat Am. Ind.Hyg. Assoc J (1984) 45(4)
Hering, Susanne V., Aerosol Science and Technology (1987) 7:257-274
Marple, V. A., Benjamin Y.H. Liu and K.T. Whitly Aerosol Sci. (1974) Vol 5, 1-16
May, K.R. (1945) J. Scient. Instrum. 22.187
May, K.R. (1975) J. Aerosol Sci. 6, 413
Mercer, T.T. and Stafford R.G. (1969) Ann. Occup. Hyg. 12.41
Mercer, T. T., Tillery, M.I. and Newton, G.J. (1970) J. Aersol Sci. 1,9
Newton, G. J., O.G. Raabe and B.V. Mokler Aerosol Sci. (1977) Vol. 8.339 to 347
Woods, D.C. and Chuan R.L. (1983) Geophys. Res. Lett. 10:1041-1044

URC97051

$\overline{5}51/08$

AIRCRAFT PITCH CONTROL WITH FIXED ORDER LQ COMPENSATORS

James Green* CR. Ashokkumar† A. Homaifar‡
NASA Center of Research Excellence
The North Carolina A & T State University
Greensboro, NC 27411

ABSTRACT

This paper considers a given set of fixed order compensators for aircraft pitch control problem. By augmenting compensator variables to the original state equations of the aircraft, a new dynamic model is considered to seek a LQ controller. While the fixed order compensators can achieve a set of desired poles in a specified region, LQ formulation provides the inherent robustness properties. The time response for ride quality is significantly improved with a set of dynamic compensators.

1. Introduction:

While designing a feedback control, ride and handling qualities are major performance objectives in aircraft control problems. Such objectives are normally achieved by closed closed loop pole assignment [1]. Preserving these closed loop poles (within the desired regions) in the presence of perturbations is another requirement [2]. LQ problems have inherent stability margins to tolerate unstructured uncertainties. LQ design techniques with regional pole constraints have been studied extensively in the literature see [3], and its references]. Similar approach, but with dynamic compensators, have been investigated for automotive applications [4]. The compensators given in [5] for aircraft control problem are considered in LQ problem setting. The objective of this approach is to improve aircraft ride quality defined in [1].

II. Problem Formulation:

An aircraft model in pitch plane [with normal acceleration (n_z), pitch rate (q) and elevator deflection (δ_e) as state variables and command input (δ_c) as control variable], is given by [2]:

$$\dot{x} = \underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & -14 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} n_z \\ q \\ \delta_e \end{bmatrix}}_{x(t)} + \underbrace{\begin{bmatrix} b_1 \\ 0 \\ 14 \end{bmatrix}}_{b} u \tag{1}$$

It is well known that the control law

$$u(t) = -R^{-1}b' Px(t) + r(t) \tag{2}$$

minimizes the performance index

$$J = \int_0^\infty \{x'Qx + u'Ru\}dt \tag{3}$$

and satisfies the algebraic riccati equation

$$A'P - PbR^{-1}b'P + PA + Q = 0 \tag{4}$$

Selection of weighting matrices to achieve a controller in equation 2 for exact pole assignment has been extensively investigated in reference [3]. Suppose, we choose a set of dynamic compensators given in [5] for the control law structure [2] (see Figure-1), then the state equations for the compensators are:

* [Undergraduate Student, Dept. of Electrical Engineering.
† Post Doctoral Research Associate.
‡ Associate Professor, Dept of Electrical Engineering.

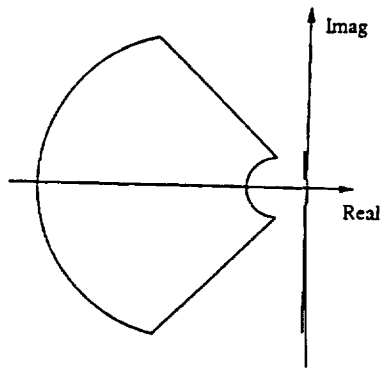


Figure 2: Regional Constraints for Aircraft

II. Simulation Results:

For F-4 aircraft model at Mach= 1..5, Altitude =35,000ft, the system dynamic matrices are given by:

$$A = \begin{bmatrix} -0.5162 & 26.96 & 178.9 \\ -0.6896 & -1.225 & -30.38 \\ 0 & 0 & -14 \end{bmatrix}$$

$$b = \begin{bmatrix} -175.6 \\ 0 \\ -14 \end{bmatrix}$$

The matrices $\bar{A}$ and $\bar{b}$ for the sate vector $\bar{x}(t) == [x(t), z_1(t), z_2(t)]'$ are

$$\bar{A} = \begin{bmatrix} A & 0 & 0 \\ A_1 & -\tau_1 & 0 \\ A_2 & 0 & -\tau_3 \end{bmatrix}$$

$$\bar{b} = \begin{bmatrix} -175.6 \\ 0 \\ -14 \\ 0 \\ 0 \end{bmatrix} \quad \text{where,}$$

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

$$A_2 = [\ -0.6896 \ \ (-1,225 + \tau_2) \ \ -30.38\ ]$$

At this flight condition, the short period damping (ζ_{sp}) and frequency (ω_{sp}) requirements are:

$$0.35 < \zeta_{sp} \leq 1.3 \tag{12}$$

and

$$3.29 < \omega_{sp} \leq 11.8 \tag{13}$$

In complex plane, these constraints impose regional pole constraints shown in Figure 2.

Table- 1

Design Variables	ω_{sp}	ζ_{sp}
Q. R	4.5078	0.4789
$\bar{Q}$, $\bar{R}$	6.44.58	0.5316
Desired	[3.29, 11.8]	[0.35, 1.3]

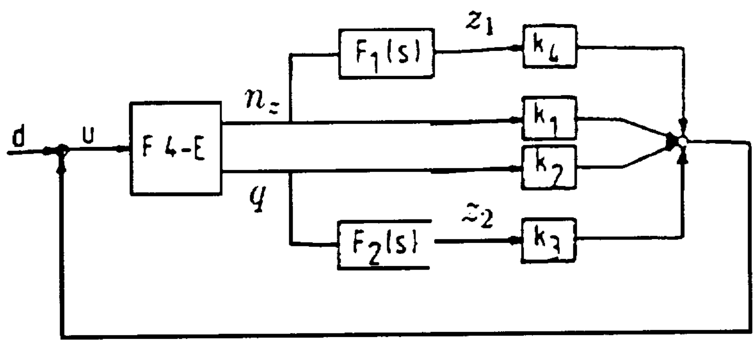


Figure 1: Control Law Structure [Ref 2] with Filters $F_1(s) = \frac{1}{s+\tau_1}$ and $F_2(s) = \frac{s+\tau_2}{s+\tau_3}$

$$\dot{z}_1 = -\tau_1 z_1 + n_z \tag{5}$$

$$\dot{z}_2 - \dot{q} = -\tau_3 z_2 + \tau_2 q \tag{6}$$

From the aircraft dynamical equations 1, substituting for $\dot{q}$, we have

$$\dot{z}_2 = a_{21} n_z + (a_{22} + \tau_2) q + a_{23} \delta_e - \tau_3 z_2 \tag{7}$$

For the new state vector z(t),

$$\bar{x}(t) = [x(t), z_1(t), z_2(t)]'$$

the state space equations become,

$$\dot{\bar{x}}(t) = \bar{A}\bar{x}(t) + \bar{b}\bar{u}(t) \tag{8}$$

where,

$$\bar{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ 0 & 0 & -14 & 0 & 0 \\ 1 & 0 & 0 & -\tau_1 & 0 \\ a_{21} & (a_{22}+\tau_2) & a_{23} & 0 & -\tau_3 \end{bmatrix}$$

$$\bar{b} = \begin{bmatrix} b_1 \\ 0 \\ 14 \\ 0 \\ 0 \end{bmatrix}$$

It can be verified that for these dynamic compensators, the system in equation 8 is completely controllable. Thus the control law

$$\bar{u}(t) = -\bar{R}^{-1}\bar{b}'\bar{P}\bar{x}(t) + r(t) \tag{9}$$

minimizes the performance index

$$\bar{J} = \int_0^\infty \{\bar{x}'\bar{Q}\bar{x} + \bar{u}'\bar{R}\bar{u}\}dt \tag{10}$$

and satisfies the algebraic riccati equation

$$\bar{A}'\bar{P} - \bar{P}\bar{b}\bar{R}^{-1}\bar{b}'\bar{P} + \bar{P}\bar{A} + \bar{Q} = 0 \tag{11}$$

With the above formulations, we shall now present the closed loop eigenvalues for various values of the design parameters. The design parameters for J are obviously the weighting matrices Q and R. However, note that the performance index $\bar{J}$ is significantly influenced by the other design parameters τ_1, τ_2, and τ_3, in addition to Q and $\bar{R}$. The next section presents the simulation results.

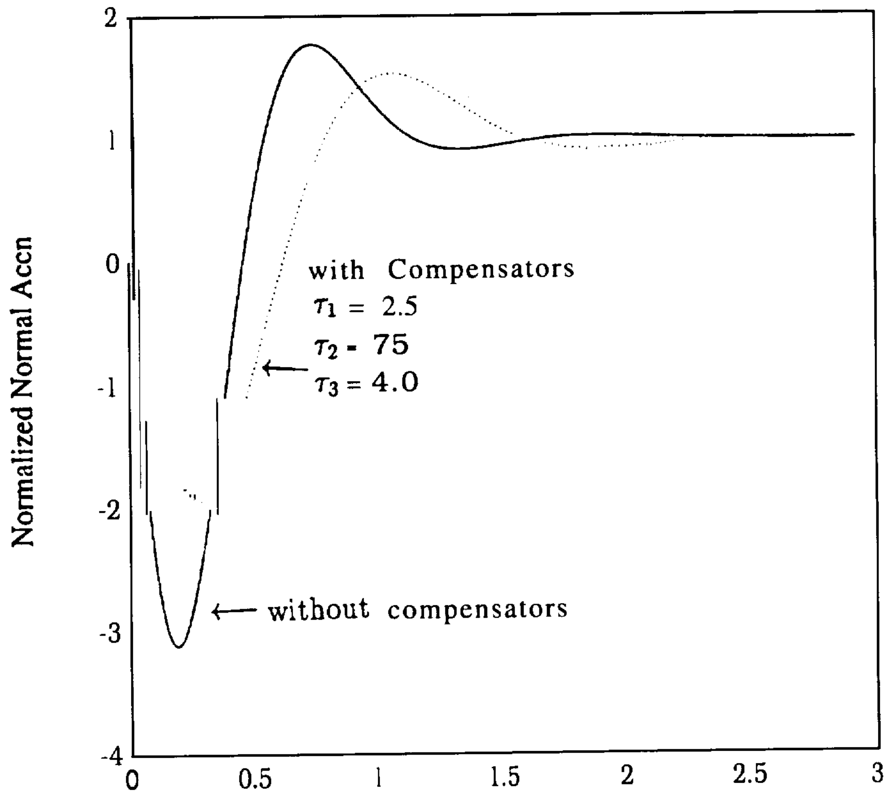


Figure 3: Time response pelts due to step input

The weighting matrices Q = $1_3$and R. = 10^4as well as the weighting matrices $\bar{Q}$ = $1_3$and $\bar{R}$ = 10^4provide the acceptable closed loop poles [see Table 1].

However, what needs to be observed is the time response plots (due to step input) shown in Figure 3. We observe that the normal acceleration at the sensor location is nonminimal. Moreover, the peak accelerations are significantly reduced with dynamic compensators (about 50%).

Acknowledgements:

This work is partially supported by grant from the NASA Center of Research of Excellence at NC A&T State University under grant # NAGW-2924. The authors wish to thank the NASA-CORE administration.

References:

1. "Flying Qualities of Piloted Airplanes," *MIL-F-8785B(ASG)*, Aug 7, 1969

2. S.N. Franklin and J. Ackerman, "Robust Flight Control:A Design Example," *Journal of Guidance, Control, and Dynamics, v 4,* n *6, 1981,* pp 597.

3. Y. Ochi and K. Kanai, "Pole Placement in Optimal Regulator by Continuous Pole-S hefting," Journal *of* Guidance, *Control, and Dynamics, v 18,* n 6, 1995, pp 1253-1258

4. H. Peng and M. Tomizuka, "Preview Control for Vehicle Lateral Guidance in Highway Automation: Journal *of Dynamic systems, Measurement and Control, v* 115, 1993, pp 679-686

5. CR. Ashokkumar,"Robust Optimal Compensators with Tight Control Philosophy: Rep. GCD/CRA/2, NASA Center of Research Excellence, North Carolina A & T State University, Greensboro, NC

URC97052

Spectroscopy of the C_2N_2 C $^1B_u \leftarrow$ X $^1\Sigma_g^+$ Transition

Joshua B. Halpern and Yuhui Huang
Department of Chemistry
Howard University
Washington, DC 20059

1. Introduction

Ethanedinitrile (C_2N_2) is an important photochemically active species in Titan's atmosphere [1,2]. It may be found in comets, where it would be a source of CN radicals whose emission is prominent in the cometary tail [3]. Moreover, this relatively simple, molecule has become an exemplary system for photochemical studies [4-13].

The first ethanedinitrile (C_2N_2) vacuum ultraviolet (VUV) absorption spectra were measured by Price and Walsh [14]. The most striking feature was a strong absorption system between 145 and 170 nm which, at least at the red end, had fairly sharp vibrational structure. Twenty years later, Bell, et al assigned this system as C $^1\Pi_u \leftarrow$ X $^1\Sigma_g^+$ [15]. The origin was given as 60,420 cm-'. The strength of the transition showed that it was electronically allowed, which for a linear-linear transition with an X$^1\Sigma_g^+$ ground state limits the upper electronic state to either $^1\Sigma_u^-$ or $^1\Pi_u$. The complicated vibrational band pattern at the red end of the spectrum spoke against $^1\Sigma_u^+ \leftarrow {}^1\Sigma_g^+$. Several three and four member progressions can be seen with about 2050 cm-1 between the bands. Molecular absorption coefficients in the VUV were measured by Connors, *et al. [16]*, West [17], and Nuth and Glicker [18].

Bell carried out *ab initio* Hartree-Fock (RHF) calculations with a double-zeta basis of contracted gaussian functions for 40 excited states of different configurations and symmetries of ethanedinitrile and four excited states of its cation $C_2N_2^+$ [19]. The assignments from ultraviolet absorption and photoelectron spectra were checked against calculated excitation energies. Dateo, Dupuis and Lester did an *ab initio* multiconfiguration Hartree-Fock calculation for the X $^1\Sigma_g^+$, a $^3\Sigma_u^+$, b$^3\Delta_u$, A $^1\Sigma_u^-$, B$^1\Delta_u$ and C$^1\Pi_u$ states [20,21]. Equilibrium geometries, vibrational frequencies, excitation and bond dissociation energies were calculated and were in rough agreement with experimental results. Table I lists the calculated properties of the X and C states. The n $\rightarrow \pi$ character of the C $^1\Pi_u \leftarrow$ X $^1\Sigma_g^+$ transition in its equilibrium configuration was confirmed. Most importantly, they predicted that the equilibrium position of the C state was very slightly trans-bent (170'3- 174°) and the cis bending frequency decreases from 233 cm^{-1} (206 cm^{-1} from the calculation) to 85 cm^{1}in the C state.

2. Experimental

Absorption spectra of ethanedinitrile were measured on two different single beam spectrometers. The first was a 0.3 in McPherson model 218 vacuum-UV monochromator operated in the second order (0.03 nm bandwidth) with a 1() watt Hellma deuterium lamp and a 2400 grooves/mm UV grating. The current output of a solar blind photomultiplier tube (EMR 541-F) attached to the sample cell was converted to voltage by a Keithley 412 logarithmic picoammeter and recorded in a computer data acquisition system. The temperature of a 158 mm long steel cell sealed with LiF windows could be varied between -78 and 100 'C by surrounding the cell body with dry ice saturated in methanol or using electrical heating tape.

The second spectrometer was a 1 m McPherson model 225 vacuum-UV monochromator with a 30 watt Hamamatsu deuterium lamp and a 1200 lines/mm grating at NIST [22]. The scan gear box was connected to a computer controlled stepping motor and the signal was sensed by a side window EMI Vacuum-UV photomultiplier tube. Baselines were first recorded in the computer, then absorption spectra were measured when the ethanedinitrile gas was introduced into the 18.'7 mm *long* sample cell. The resolution of spectra taken in this instrument was about 0.05 nm. Ethanedinitrile from Matheson was purified by freeze-thaw cycling.

Finally, the spectra from Reference 18 were provided to us as a computer file by Drs. Nuth and Glicker.

3.0 Results

Figure l shows the absorption spectrum associated with the transition to the C state of ethanedinitrile. Three large clumps of lines, separated from each other by about 2050 cm^{-1}, are found in this spectrum. Figure 2 shows the spectrum between 58,400 and 61,600 cm-1 measured with an instrumental resolution of 20 cm^{-1}. Figure 3 shows the expanded and differential spectra in the hot band region below 61,600 cm^{-1} where the latter is the difference between the sample's absorption cross-section at 22 and -78 'C. Table 11 lists the measured band positions and peak absorption cross-sections.

Even though the origin of the excited state is about 12,800 cm-1 above the ethanedinitrile dissociation limit, the first group of lines between 61,500 and 58,800 cm^{-1} still has a well resolved vibrational-structure [14- 18]. The narrowest band is about 20 cm^{-1} wide, which can be completely accounted for by instrumental broadening. Absorption bands are strongly predissociated above 62,000 cm^{-1} and thus are much broader. There is no clear trend of increasing bandwidth as a function of excitation energy, but rather a sudden change at about 62,000 cm^{-1}, where some barrier to dissociation is exceeded.

4.0 Spectral Analysis

Despite the low spectral resolution, substantial progress can be made in analysis of the transition using spectra taken at different temperatures and Dateo, et al.'s *ab initio* calculation. Since the spectrum is predissociative higher resolution would not help, especially below 62,000 cm^{-1} where the predissociation rate increases sharply. There are undoubtedly many vibrational and electronic perturbations which are hidden from view by the both the low resolution of the spectrum and predissociative broadening. Thus, great care should be taken in using any spectroscopic constants from this analysis. It does, however, provide an excellent qualitative picture.

Cold and hot bands could be distinguished from the differential spectrum (AT= 100 'C). The first cold band was found at 59,842 cm-', which we assign as the band transition origin. This is about 560 cm-' lower than the value given by Bell et al [15]. Refering to Table II and Figure 3, and given the known ground state vibrational frequencies, hot bands can be immediately assigned corresponding to excitation of the $X^1\Sigma_g^+$ $\nu_2'' = 1$, $\nu_4'' = 1$ and 2 and the $\nu_5'' = 2$ levels to the excited state vibrationless level. The splittings between Σ and Π 42 and 52 levels are 14.7 [23] and 2.63 cm^{-1} [24] respectively, less than the resolution of the monochromator, and they are not resolved.

Identification of the hot and cold bands makes it much easier to determine the upper electronic state symmetry. For an electronically allowed linear-linear transition, only bands linking totally symmetric vibrational levels will be strong. Unit changes in vibrational quantum numbers between the ground and excited state would be restricted to two totally symetric vibrational modes, $\nu_1(\sigma_g)$ and ν_2 (erg). For a fully allowed transition the change between the ground and excited state vibrational quantum numbers would have to be even for the ungerade modes, $\nu_3(\sigma_u)$, $\nu_4(\pi_g)$ and $\nu_5(\pi_u)$. The bending potential energy minimum would remain at 180°, so by a Franck-Condon argument, allowed 4_0^{2n} or 5_0^{2n} progressions would be very weak. A similar argument can be made for 3_0^{2n} stretching mode [25].

The *ab initio* calculation provides estimates of both ν_1' and ν_2'. The $X^1\Sigma_g^+$ ν_1'' symmetric CN **stretching** frequency is 2330 cm^{-1}. Then $\rightarrow\pi$ electronic transition will decrease this vibrational frequency, according to Dateo, *et al.*, to 2145 cm-i, They further predict that the $\nu_2(\sigma_g)$ symmetrical CC stretch will increase from 890 cm-' to 989 cm-' in the excited state since the CC bond strengthens and the CN bonds acquire more of a double bond character in the excited state. If bond distances did not change substantially between the ground and excited state, the 0_0^0 band would be the strongest. This is clearly not the case, again, in agreement with the *ah initio* calculation. For a substantial change in internuclear distances, with a linear excited state, the cold bands would consist of simple progressions in ν_1' (erg) and ν_2' (erg). *This is **not** observed.*

The presence of a 4_1^0 hot band offers another clue that the upper state is not linear. In a I inear to trans-bent transition, unit changes in vibrational quantum numbers would also be found for vibrational bands involving the trans-bending ground state $\nu_4''(\pi_g)$ and the excited state $\nu_3'(a_g)$ modes. For a non linear upper state which retained its center of inversion, transitions involving ungerade vibrations in both upper and lower electronic states wou Id still be allowed only if the net change in vibrational quantum numbers was even. In the ground state these are the $\nu_3''(\sigma_u)$ antisymmetric stretch and the $\nu_5''(\pi_u)$ twofold degenerate cis-bending modes. In the electronically excited state the ungerade vibrations are the $\nu_5'(b_u)$ antisymmetric stretch, the ν_4' (au) torsion and the $\nu_6'(b_u)$ in plane bend.

With this in mind, four of the cold bands shown in Figure 2 can be assigned to a simple progression in the excited state trans-bending mode $\nu_3'(a_g)$. This agrees with Dateo, et al.s, calculation that the C state equilibrium configuration is trans-bent. The n - π electronic transition is the result of promotion of a $5\sigma_g$ electron to the $2\pi_u$ antibonding orbital. In the trans-bent C_{2h} symmetry, the $5\sigma_g(a_g)$ orbital and the Renner-Teller split all component of the $2\pi_u$ antibonding orbital do not vary much with bending angle. However, the b_u component of the $2\pi_u$ orbital is strongly stabilized by bending. Therefore, we conclude that the upper state is C^1B_u.

The absence of strong progressions between the ground state 0_0 level and the ungerade ν_4' and/or ν_6' = 1,3 shows that the excited state retains its center of inversion with a trans-bent shape and belongs to the molecular point group C_{2h}. This is further confirmed by the absence of strong hot bands connecting the 5_1 and 53 levels with the excited electronic state vibration less level. We do, however, observe two small lines corresponding to 5,0 hot band and excitation of one quanta in either the ν_4' and/or ν_6' level, which we will refer to below as ν_u. Thus, strictly speaking we must refer to the C state as having only near trans-bent symmetry.

Most of the cold bands in Figures 2 and 3 are the first members short red degraded progressions. The separation between bands is about 95 cm^{-1}, with a variation that is less that the resolution of the monochromator. The differential spectrum shows that these are hot bands, the relatively high aborption cross-section suggests that they are associated with thermal excitation of the $X^1\Sigma_g^+$ lowest frequency vibrational mode, $\nu_5''(\pi_u)$. Transitions will be strongest connecting these levels with excited vibrational levels of the $C\ ^1B_u$ $\nu_4'(a_u)$ torsion and the $\nu_6'(b_u)$ in plane bending modes. The spectra cannot differentiate between these two modes. Since the 51 level is 233 cm-' above the ground vibrationless state, and the separation between the 0_0^0 and first hot band to the low frequency side is 95 cm"', the first excited level of the ungerade cis-bending vibration must lie at about 138 cm-i above the $C\ ^1B_u$

state origin. This agrees qualitatively with Dateo, et al.s prediction that the C^1B_u state cis-bending frequency would be substantially less than that in the ground state. Progressions as many as four members connecting 5_n with vu' = n. account for much of the structure below 62,000 cm^{-1} and can be assigned in a straightforward manner.

Inspection of the region to the blue of the electronic origin shows that there are absorption at frequencies corresponding to excitation of two quanta in vu' as well as a hot band connecting 51 with vu' = 3. These bands are relatively intense because of the mis-match between the ground state cis-bending vibration and the cis-bending and out of plane torsion in the C^1B_u state. Several other bands in the spectrum can be assigned as combination bands involving a change of two quanta in the cis-bending/tors ional modes the largest of which at involves excitation of one quantum in the excited state ν_3' mode and two quanta in the ν_u' mode..

One can see a shoulder between the 3_0^1 band and the associated hot band immediately to its red side which would correspond to a 2_0^1 band at about 960 cm-'. Values in Table II are from a global fit, so they differ a bit.

In order to assign bands involving excitation of one or more quanta in the symmetric CN stretch, ν_1' we can start from the assigned cold bands below 62,000 cm-1. Figure 1 and Table II give the assignments. The spacing between bands is about 2030 cm-' $\pm$ 20 cm^{-1}. One cannot assign all of the largest bands above 62,000 cm^{-1} using a ten percent larger or smaller value for ν_1'. As the amount of CN symmetric stretch excitation increases, the upper state becomes increasingly bent. This is shown by the increased relative intensities of the bands with two and three quanta in the ν_3' stretching mode when there are one or more quanta in the ν_1' mode. The last progression is a further surprise. It starts about 1715 cm-' from the band origin, and the first band belonging to this progression: is very weak, however, the highest wavelength band is relatively strong. We would tentatively identify this as the $1_0{}^n5^1$ progression, involving excitation of one quanta in the asymmetric stretch. Only high on the potential energy surface, where the wavefunction has more cis character, are transitions involving a unit change in the $\nu_5'(b_u)$ antisymmetric stretch quantum number significant.

We see that both the 0_0^0 and 1_0^1 bands are weak, especially compared to the transitions involving excitation of one or more quanta in ν_3' Therefore, even if transitions to the calculated linear component of the C state are present in the spectra of Figures 2-4, they do not account for much if any of the absorption.

Final ly, it would be useful to restate the limitations of our analysis. We have ignored perturbations because of the low instrumental resolution and the predissociatively broadened nature of the spectrum makes it impossible for us to locate any. The calculation of the cis-bending frequencies in particular neglects the mutual perturbations of the ν_4' and ν_6' modes, lumping them together as an antisymmetric bending vibration. Predissociation broadening and low resolution make it impossible to find x_{nm} enharmonic constants.

4.0 Conclusion

The $C^1\Pi_u \leftarrow X^1\Sigma_g^+$ absorption spectra of ethanedinitrile was measured, and partially assigned in the 145-230 nm region. Values were determined for the band origin, 59842 cm^{-1}, the ν_1' frequency of 2030 cm^{-1}, the ν_2' frequency of 945 cm-1, the ν_3' frequency, 533 cm^{-1}, ν_5' frequency of 133 and the ν_5' frequency of 1715 cm-i.

In closing, we must acknowledge our debt to the many people who provided spectra to us, or helped us take data on our instruments. These include our thanks to Dr. Joseph Nuth and Sol Glicker of NASA Goddard Space Flight Center for their C_2N_2 data file, and Dr. Walter Braun, and Allan Laufer for helping us measure spectra on their VUV monochromator and spectrometer.

5.0 References

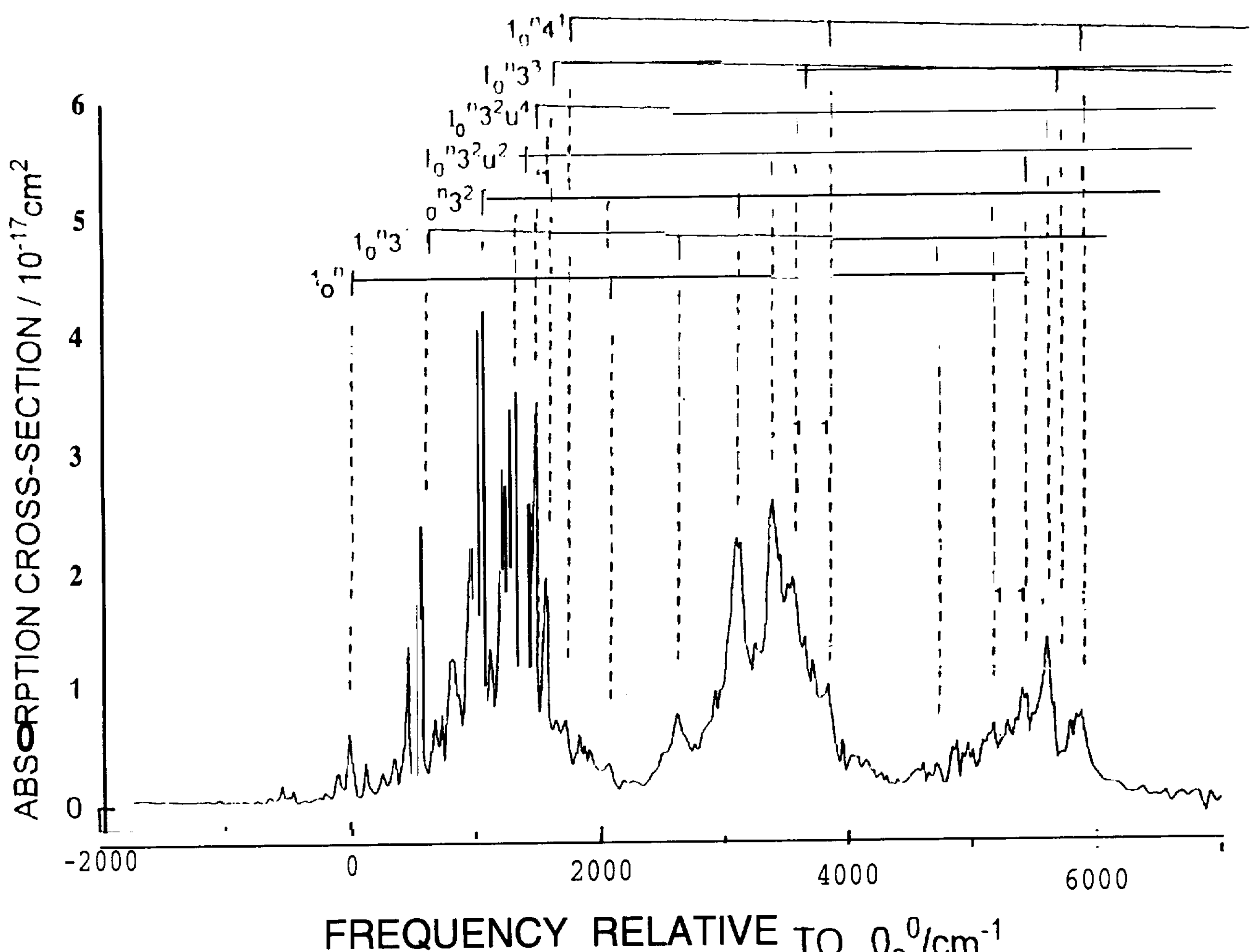
1. Y.L.Yung, M. Allen and J.P. Pinto, *Astrophys. J. Suppl.*, **55** (1984) *465.*
2. A. Coustenis, B. Bezard and Gautier, D. *Icarus, 80(1989) 54.*
3. D. Brockelee-Morvan and J. Crovisier, *Astron. Astrophys.*, 151, *(1985) 90.*
4. G.E. Miller, W.M. Jackson and J.B. Halpern, J. *Chem. Phys.*, 71 (1979) 4625.
5. M. R. Taherian and T. G. Slanger, *J. Chem. Phys.* 81 (1984)3814
6. J. B. Halpern and W. M. Jackson, J. *Phys. Chem., 86,* (1982) 973.
7. R. Lu, J. B. Halpern and W. M. Jackson, J. *Phys. Chem.,, 88, (1984) 3419.*
8. D. Eres, M. Gurnick and J. D. McDonald, *.I. Chem. Phys.*, **8***1, (1 984) 5552.*
9. H. Lin, E. A. J. Wannemacher and W. M. Jackson, *Chem. Phys. Letts., 152,* *(198*8) 477.
10. E. A. J. Wannemacher, H. Lin and W. M. Jackson, *J. Phys. Chem., 94, (1990) 6608.*
11. Y. Huang, S. A. Barts and J. B. Halpern, *J. Phys. Chem.* 96, (1991) 425.
12. M. Wu and G.E. Hall, J. *Photochem. Photobiol A: Chem, 80, (1994) 45.*
13. S. W. North and G.E. Hall, submitted to *J. Chem. Phys.*
14. S. Bell, G. J. Cartwright, G, Fish, D. O. () 'Hare, R. K. Ritchie, A. D. Walsh and P. A. Warsop, *J. Mol. Spectrosc.*, 30(1969) 162.
15. W'. C. Price and A. D. Walsh, *Trans. Faraday Soc. 41 (1945)381.*
16. R. E. Connors, J. L. Roebber and K. Weiss, *J. Chem. Phys. 60(19.74)5011* .

17. G. A. West, Ph. D. Dissertation, University of Wisconsin-Madison, 1975.
18. J. A. Nuth and S. Glicker, *J. Quant. Spectrosc. Radiat. Transfer* 28 (1982) 223.
19. S. Bell, *Chem. Phys. Lett.* 67(1979) 498.
20. C. E. Dateo, M. Dupuis and W. A. Lester, Jr. *J. Chem. Phys.* 83 (1985) 265.
21. C. E. Dateo, personal communication.
22.. W. Braun, personal communication.
23. A.G. Maki, *J. Chem. Phys.*, 43 (1965) 3193.
24. G. Herzberg, Molecular Spectra and Molecular Structure, III. Electronic Spectra of Polyatomic Molecules, Van Nostrand, NY.

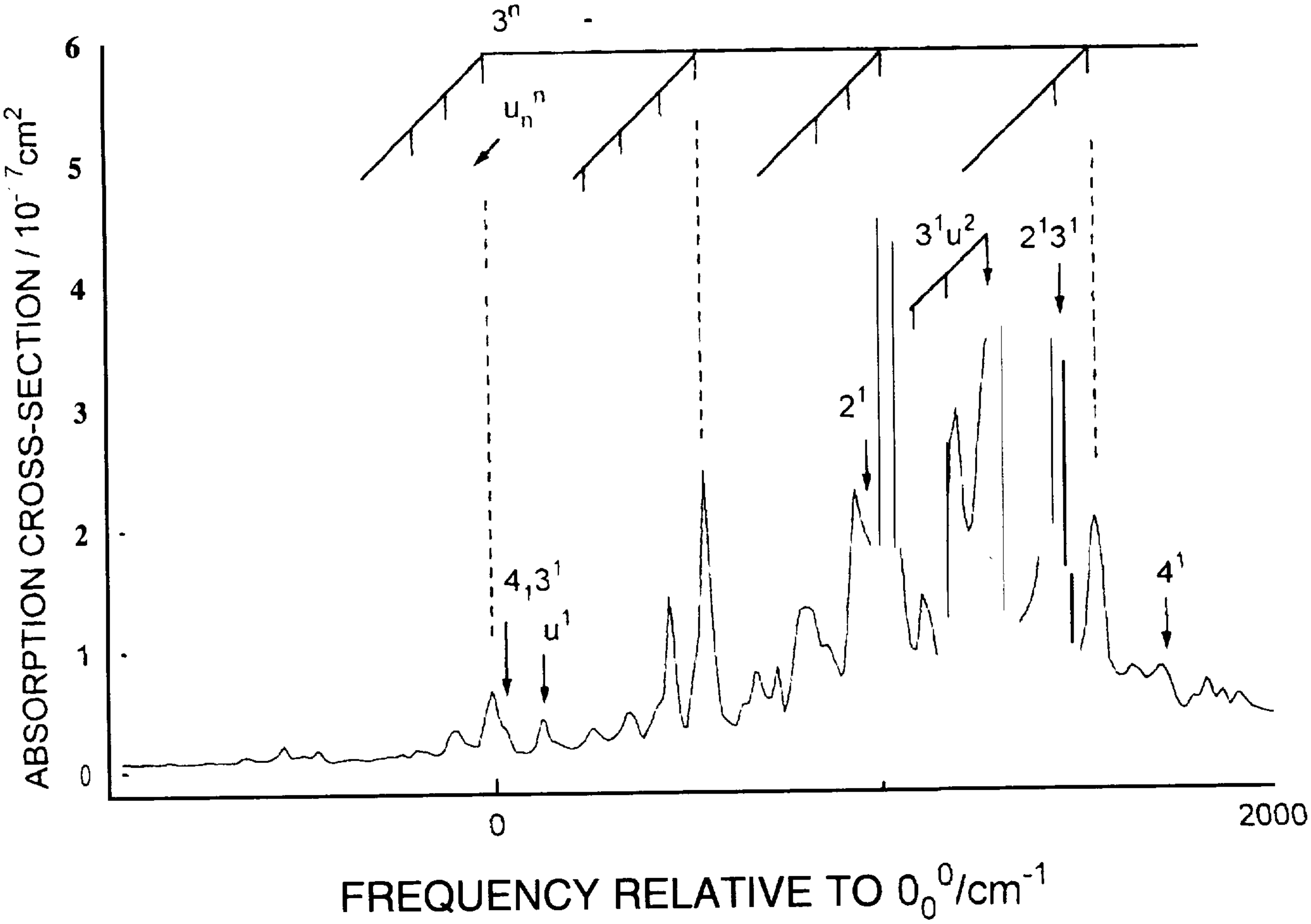
TABLE I

Comparison of C_2N_2 state energies, bond lengths and vibrational frequencies from observation and calculation.

	Ref.	T_e/eV	ω_1/cm^{-1}	ω_2/cm^{-1}	ω_3/cm^{-1}	ω_4/cm^{-1}	ω_5/cm^{-1}	R(CC)/nm	R(CN)/nm	B/cm^{-1}
$X\ ^1\Sigma_g^+$										
Obs.		0.00	2375	862	2191	507	234	0.1389	0.1154	0.1571
Calc	20	0.00	2482	895	2283	416	213	,0.1429	0.1201	
Calc	21	0.00	2457	887	2271	427	230	().1401	0.1170	0.1540
$C\ ^1B_u$			ω_1/cm^{-1}	ω_2/cm^{-1}	ω_5/cm^{-1}	ω_3/cm^{-1}	ω_u/cm^{-1}			
Obs	TW	7.3	2030	945	1715	533	133			
Calc[a]	20	8.2	2145	989	1789	698	85	0.1315	0.1232	



Nuth Glicker	This Work	ν_1'	ν_2'	ν_3'	ν_0'	ν_5'	ν_2''	ν_4''	ν_5''	Nuth Glicker	This Work	Calc	Difference
Absolute	Absolute									Relative	Relative	Relative	
cm-q	cm^{-1}									cm-l	cm-l	cm-l	
	58607				2			2	2:		-1231	-1213	-18
	58709				1			2	1		-1129	-1114:	-15
	58799										-1039	-1013	-26
	58885						4				*-953*	*-932*	*-21*
	58995						1				*-843*	*-845*	*2*
	59101				*2*			1			*-737*	*-703*	*-34*
	59194				*1*			1	*1*		*-644,*	*-604*	*-40*
	59296							1'			-542	-503	-39
	59349								2		-489	-469	-20
	59385								2'		-453	-465	12
	59539				3				3		-299	-304	5
	59602								1		-236	-234,	-2
59661	59644				2				2	-184	-194	-200	11
59745	59745				1				1	-100	-93	-101	5
59845	59838	0	o	0	0		0	0	0	0	0	0	0
59878	59869			l i				1		33	31	38	-6
59980	59968				1					'135	130	1331	0
60017	59996				3				1	172'	158	165	0
60097	60097:			1	3				3	252	259	236	19
60117,					2				0	272		266	6
60197	60191			1	2,				2	352	353	340	12
60285	60299			1	1				1	440	461	'440	11
60361	6 0 3 6 5			2				1		516	527	532	-11
60384	60392			1						539	554	541	6
60482	60490			2	4				4	637	652	635	10
60577	60579			2	3				3	732	741	731	6
60629	60642:			1	2					784	804	806	-12
60694	60706			2	2				2	849	868	835	24
60777	60788			2	1				1	932	950	934	7
60804	60815,		1							959	977	968	0
60857	60866			2						1012	1028	1035	-15
60946	60958			2	4				2	1101	1120	1.101	10
61028	61047			2	3				1	1183	1209	1200	-4
61132	61140			2	2					1287	1302	1301	-6
61297	61315		1	1						1452	1477	1487'	-22
61388	61388			3	0					1543!	1 550/	1547	o
61545,	61567					1				1700	1729	1715	-1
61766	61763			4						1921	1925	1908	15
61893	61916	1								2048	2078	2026	37
62421	62415	1		1						2576	2577	2566	10
62918"	62900	1		2						3073	3062	3061	7
63180	63185	1		2	2					3335	3347	3326	15
63349	63346	1		3						3504	3508	3520	-14
63646 !	63631	1				1				3801	3793	3741	56
64934	649391	2		2!						5089	5101	5086	9
65157	65177	2,		2	2					5312	5339	5352	-26
65351	65355	2		3						5506	5517	5546	-34
65607:		2				1				5762		5766	-4



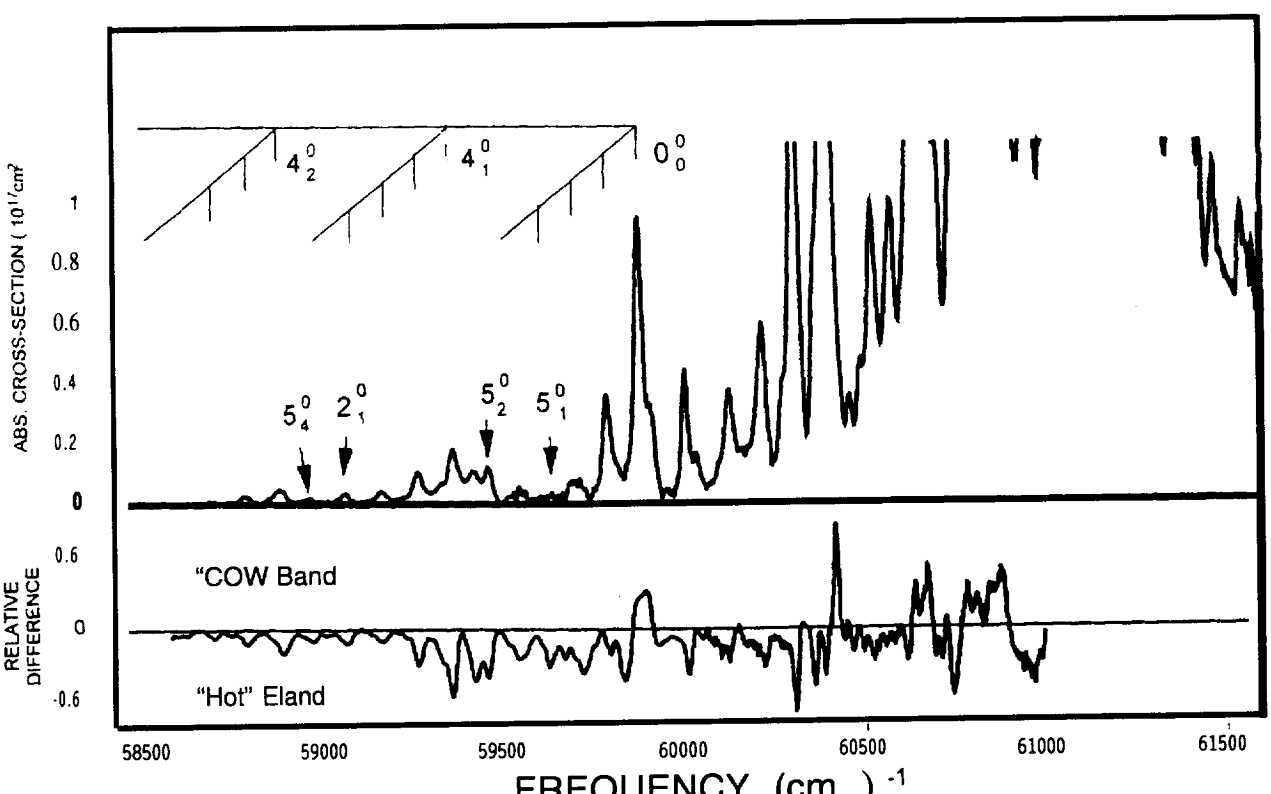


Figure 3 Expanded absorption spectrum in the 58400 - 61600 cm⁻¹ region

URC97053

INTEGRATED DISTRIBUTED INTELLIGENT AGENTS (IDIA) IN INDUSTRY

61

Mohamed L. Hambaba
Associate Professor; Computer, Information, and System Engineering (CISE) Dept.,
College of Engineering, San Jose State University, San Jose, CA

1.0 Introduction

The IDIA architecture is based on interacting intelligent agents; multi-agent architecture. It is programs that encapsulates engineering tools. In an engineering experiment, each agent models different aspects of a machine tool and reasons about them from the standpoint of a different engineering discipline. Essentially, IDIA can be envisioned as a network of agents or knowledge modules that communicate through distributed client/server based on OMG's CORBA standard, [1].

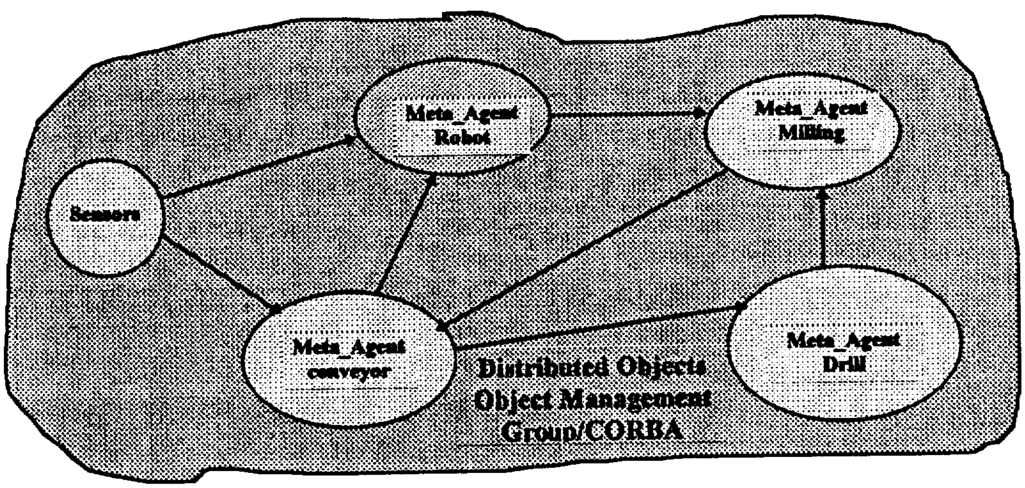


Figure 1. Mets-Agent automata structure under OMG/CORBA information sharing.

In our instantiation of this vision, illustrated in Figure (1), these roles are played, respectively, by meta agent network and the object management group CORBA for message contents and information sharing. Meta agents interact with the information resources through the model via

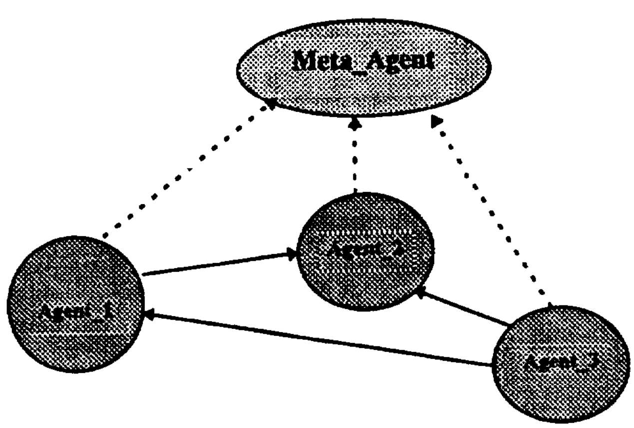


Figure 2. Mets-Agent structure.

high-level service protocol, CORBA, that insulate them from details such as where information resides. Each meta agent is another intelligent network composed of agents or concepts for a particular function within the meta agent. Figure (2) illustrates the meta agent architecture.

Although concurrent engineering is almost universally advocated today, it is hard to execute when large multidisciplinary projects are involved. To illustrate some of the issues, consider the product design team, at any instant, the team members maybe working at different levels of details, each employing his or her own representation of physical objects, engineering models, and knowledge. despite their differences in perspective, the specialists share considerable information. We require a system to reason about the level of granularities. This is called categorization.

Within the manufacturing world [2-5], the design process is very dynamic. Some design change can be introduced anytime anywhere in the design. This will slow down the design model and the decision making is suboptimal. To overcome these difficult problems, intelligent agents must have the ability to dynamically evolve in space and time to respond effectively to the design changes or unforeseen sensor inputs. One of the solution is to dynamically categorize distinctions into useful concepts and use these concepts effectively is a more competency of intelligent decision making. To approach this dynamic categorization, the utility-based method to categorization is used. It is founded on the idea that categorization is in service of action. The choice of concepts and the granularity level of concepts employed by an agent plays a very critical role in the selection of appropriate action decision.

2.0 Fuzzy Probabilistic Network for Mets-Apent

For most engineering systems, there are two important information sources: sensors which provide numerical measurements of variables, and human experts who provide linguistic instructions and description about the system.

Fuzzy probabilistic network models the meta-agent. This architecture consists of dynamic nodes which represent intelligent agents. Each dynamic node is an adaptive neuro-fuzzy system, to model an intelligent agent based on fuzzy logic theory and equipped with a feed-forward neural network architecture. It is noticeable that the logic-based approach has been already become a viable and attractive alternative in modeling complex systems. In particular, manufacturing processes have proved less amenable to an efficient quantitative analysis, because of their enormous complexity.

The agent architecture is based on fuzzy neuron graphical network. This network exhibits a strong logical structure using the neural network architecture and fuzzy logic. Within the neural network architecture, the agent will have the capability of learning, adaptation of changing environment, and knowledge storage. A strong logical structure based on@ AND and OR logic operator, $\otimes$ and $\oplus$, enable an agent to acquire easily knowledge and to reason for a given sensor observations. In general, this architecture is called fuzzy expert networks. In set theory, AND and OR operators are complete logic set to describe production rules. This structure is aimed at carrying out approximation of the logical relationships between input vectors and their associated output vectors. The fuzzy neuron network gives rise to a" so-called Sum-Of-Minterms (SOM), [6] in digital logic, Figure (3).

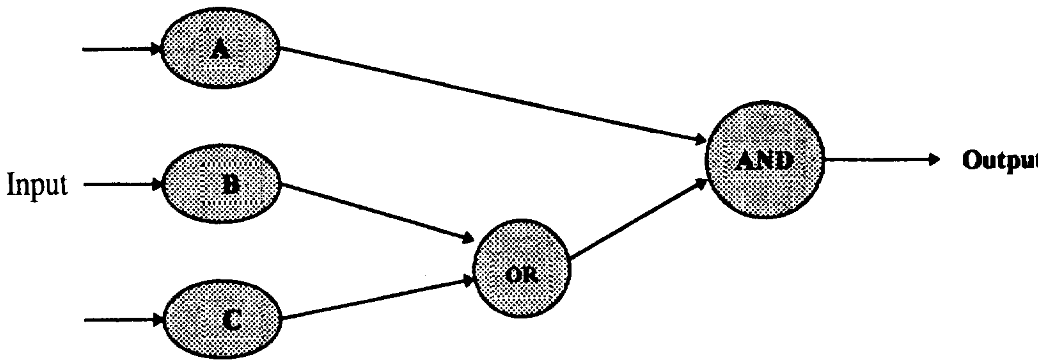


Figure 3. Agent structure: Fuzzy probabilistic network.

Let us denote the connection between the node i in the input layer and the jth neuron in the hidden layer by v_{ji}. The connection between the neuron j in the hidden layer to the neuron in the output layer is described by w_j. Thus the fuzzy neuron network equation is governed by the following expressions:

$$\text{output-hidden layer } y = \sum_{j=1}^{h} (w_j \otimes z_j), \qquad \text{hidden-input layer } z_j = \sum_{i=1}^{k} (v_{ji} \oplus x_i)$$

where x_i is the input and z_j is the output.

It is important to have an intelligent information techniques that allow flexibility in the way knowledge can be encoded, represented and discovered. Fuzzy probabilistic graphical models offer such a technique. Fuzzy probabilistic graphical models are a framework for structuring, representing and decomposing. First, let us consider the probabilistic graphical network [7], and then how we can incorporate fuzzy logic theory within the probabilistic network to handle the fuzzy qualitative aspect of the domain.

In most of system identificaton, the system has to learn the underlying probabilities, [8], and also the network structure that encodes them. If U contains only discrete variables, or if it is continuous U could be decritized using mutual information method, one often uses the *Dirichlet* distribution because it has several convenient properties; in particlar it is a recursive distribution that could be updated. It is given by

$$p(U|\zeta) = \frac{\Gamma(\sum_k N_k)}{\prod_k \Gamma(N_k)} \prod_k \theta_{x=k}^{N-1_k}$$

where $\Gamma(\cdot)$ is the gamma function. The number of network structures for a domain containing k variables or agents is more than exponential in k. Several figure-of-merit metrics or scoring rules are used to compare networks of different topology and probability distributions. Figure-of-merit metrics can be defined using the following equation:

$$\Psi(\theta,\gamma) = f(p(U|FP_s,\Theta_s,\zeta)) + g(FP_s,\Theta_s)$$

where the first component is a measure of predictive error, and the second component is a penalty for the number of parameters. This class of metric is related to the Minimum-Description-Length principle (MDL) in information theory.

How fuzzy probabilities are expressed in fuzzy probabilistic network: assume that X and Y are fuzzy random variables whose probability distributions on finite sets are described in linguistic terms. Representing $p(X)$ and $q(Y|X)$ in the form of graphs

$$p(X) = \sum_i (P_i \times P_i') \qquad \text{and } q(Y|X) = \sum_j (Q_j \times Q_j')$$

then for the joint probability distribution is as follows:

$$h(X,Y) = p(X)q(Y|X) = \sum_{i,j} (P_i \, x \, Q_j) \, x \, (P_i' * Q_i')$$

where * represents the operation of multiplication of fuzzy probabilities P_i' *and* Q_j'. It shows that this representation is decomposable which makes feasible the integration of both theories to be imbedded in a directed acyclic graph.

Research has to be done in: (1) a novel learning algorithm to link the logic-based node and probabilistic causal links, (2) incorporating a prior structure, (3) combining structure search method and evaluation during the process of automated network construction.

3.0 Fuzzy Cognitive Maps and Feed-Back Systems for Metareasoning or Control of Reasoning

The fuzzy cognitive map is a directed cognitive graph consisting of nodes with connections (edges) that describe the causal flow. The nodes are meta-agents - such as $MA_1, MA_2, \ldots, MA_K$ - and the edges indicate the degree to which MA_i causes or results from MA_j. Figure (4) shows the organization of a typical fuzzy cognitive map.

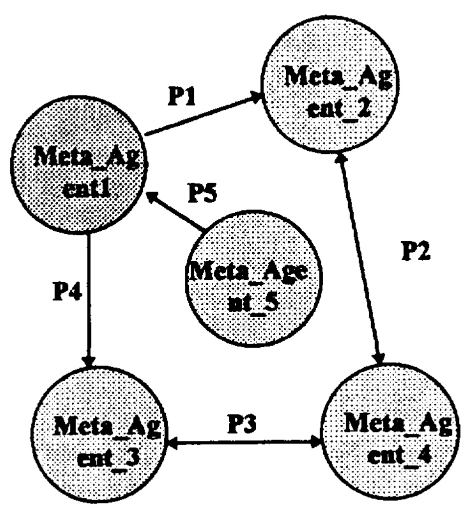


Figure 4. Typical fuzzy cognitive map.

It has been clear for some time that organization is a **powerful** concept for thinking about how to structure the interactions of problem solvers or **meta-agents**. Understanding the concept of organization and developing techniques for adaptive reorganization are pressing concerns in multi-agent architecture or distributed artificial intelligence.

Self-organization for **meta-agent (SOMA)** is proposed to allow an organization of **meta-agents** to adapts itself to changing situations. The reorganization mechanism consists two **powerful** methods:

- *Construction of temporal metastructure model*: typically changing **meta-agent** roles or **inter-meta-agents** task ordering
- *Composition and decomposition of meta-agents* (categorical reasoner or **metareasoner**): building and solving categorization decision models requires decision making about the level of abstraction. Model abstractions introduces tractability of decision making inference at the expense of decision of quality.

One has to apply scoring metrics for model constructions with respect to predictive accuracy and computational efficiency. A temporal megastructure consists of sequence of different network structure indexed by time, $\vartheta_0, \vartheta_1, \ldots, \vartheta_t$ or $\vartheta_t = \vartheta(V(t), A(t))$. A temporal arc $A_\tau(t)$ connects networks for time interval $t - \tau \le time \le t$.. For a first-order **Markov** arc $A_\tau(t)$ *is* **such** that $A_\tau(t) \in V(t-1) \times V(t)$, and second-order **Markov** arc xl.(t) includes arcs given by $AT(t) \in [V(t-1) \times V(t)] \cup [V(t-2) \times V(t)]$. Thus, the temporal **metastructure** at time step P is given by $' = *(V', A')*, where

$$V^P = \bigcup_{t=0}^{P} V(t) \text{ and } A^P = \bigcup_t A(t) \cup \bigcup_t A_\tau(t)$$

A control technique must be developed for making decisions about megastructure improvement. One must balance the expected benefits of model improvement in term of the recommended decision action with the corresponding change in computation cost. We define the net model improvement value **(MIV)** for megastructure $\vartheta(V, A)$ to be

$$MVI(\vartheta(V,A)) = EVC(\vartheta(V,A)) - EVC(\vartheta(V^0, A^{\text{``}}))$$
$$netMVI(\vartheta(V,A)) = MVI(\vartheta(V,A)) - \Delta C_c(\vartheta(V,A)) - C_g(\vartheta(V,A))$$

where $\vartheta(V^0, A^{\,\circ})$ is some reference megastructure. $\Delta C_c(\vartheta) = C_c(\vartheta) - C_c(\vartheta^0)$ is the change in computation cost, and C_g ($) is the cost of **performing** the model improvement.

4.0 Literature Cited

[1]. R. **Orfali**, Dan **Harkey**, J. Edwards, The Essential Distributed Objects Survival Guide. Wiley, 1996.

[2]. S. H. Huang, H-C Zhang, "Neural-expert hybrid approach for intelligent manufacturing: A survey", ***Computers in Industry,*** 26 (1995), pp. 107-126.

[3]. M. Hambaba *et al.*,"Intelligent Framework for Part Design", ***Journal of Intelligent & Fuzzy Systems,*** Vol. 2, issue 1, 1994, pp. 89-97,

[4]. Ingemar Huthage,"The Architecture of ALADIN: A Knowledge-Based Approach to Alloy Design", *IEEE Expert,* August 1990, pp 56-66.

[5]. S. H. Huang *et al.,* "Function approximation and Neural-Fuzzy Approach to Machining Process Selection", ***IEEE Trans. Comp. Pack. and Manuf. Tech.,*** vol. 19, No. 1, January 1996, pp. 9-18.

[6]. W. Pedrycz, "Distributed Fuzzy System Modeling," ***IEEE Trans. on Systems, Man and Cyb.,*** vol. 25, No. 5, May 1995, pp. 769-780.

[7]. E. Charniak, "Bayesian Networks Without Tears," ***AI Magazine***, 12:,50-63.

[8]. I. Good, The Estimation of Probabilities, Cambridge, Mass.: The MIT Press. 1965.

URC97054

S34/61

How difficult is it to add 1? A pedagogical example of how theory of computing may be useful

Michael Hampton

Center for Theoretical Research and its
Applications in Computer Science (TRACS)
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968
email mhampton@cs.utep.edu

Abstract

We show that while adding two generic b-bit integers requires $\approx b$ bit operations, adding 1 to an integer requires, on average, only 2 bit operations. This fact explains why the operation of adding 1 is often separately hardware supported, and why this operation is often separately described in high-level programming languages like C++.

This results shows that t heretical analysis can help in deciding which operations must be hardware supported, and thus. hopefully, will help in designing faster computers.

1 General Introduction

For NASA-oriented computer applications, it is very important to make computers faster. From the computer science viewpoint, one of the major problems of NASA research is that so much data is coming from every space mission that the existing computer systems can process, by some estimates, only about 10% of this data.

How can we *increase the* throughput? Since the main problem is that computers cannot process all the data, a natural solution is to make computers *faster*.

At first glance, making computers faster is a computer engineering problem, but computer science can also help. In principle, there are two ways to make computers faster:

- First, we can try to further miniaturize the electronics by making the chip's elements even smaller. This is desirable but extremely difficult.
- Second, even within the same technological level, we can change the computation speed by selecting different sets of computer operations to be hardware supported.
  - When an operation (e.g., addition, multiplication, etc.) is implemented in ***hardware,*** it is extremely *fast*;
  - otherwise, if this operation is ***not*** directly ***hardware*** supported, we have to implement it on the microprogramming level, as a sequence of two or more hardware supported operations, which make computations slower-.

[t is therefore desirable to be very careful in choosing which operations we want to hardware support: ideally, we should support, the operation that are used most frequently, so that the resulting slow down of not supported (less frequent) operations will have the smallest impact on the resulting computation speed.

At first glance, the problems related to both ways of increasing computer speed lie in ***computer engineering***. However, as we will show in this paper, at least for the second approach, ***computer science*** (and. especially, ***theoretical computer science)*** can be very useful.

Namely, we will show, on a simple example, that some existing ***semi-heuristic*** choices of the hardware supported operations can be explained and formally justified in t heretical computer science. This example shows that in more complicated situations, in which an optimal choice of hardware supported operation is not yet **known,** methods of theoretical computer science can he of help.

2 Formulation of the Case Problem: Adding 1 and Why It Is Import ant

Adding 1 is an important particular case of addition. One of the simplest arithmetic operations, that is directly hardware supported on all the computers, is addition of two integers.

Not all additions are born equal. If we try to trace, on the level of computer instructions, what exactly additions are performed in a typical data processing algorithm, then we will see that these additions fall into the following two categories:

- First, there are additions of data values, that come directly from the additions in the original program (written in Fortran, C, C++, or any other high-level programming language).
- Second, additions that were "hidden" in the FOR loop or in other construction of the original program. The most typical of such hidden addition is a FOR loop, in which a loop index i starts with its lower bound, and then gets increased by 1 on each iteration until it reaches its upper bound.

In other words, a reasonably frequent addition is adding 1: `i:=i+1`. This particular case of addition is so frequent that in C++, there is a special denotation for adding 1, as opposed to any other addition: the famous notation i++ that has lead to the very name of this language.

Adding 1 is implemented in hardware: why? In several processing languages, adding 1 is described as a separate operation; in several processors, it is hardware supported as a separate operation, ***in addition*** to the general addition.

This support is, usually, done on a ***semi-heuristic*** basis, in the sense that it is not justified by any precise arguments.

3 How Can We Decide Which Operations to Support: an Idea

When does it make sense to design a special hardware support for a particular case of an already hardware supported operation?

If we try to support too many operations in hardware, then the very choice of an operation will take more computation time. Thus, every additional operation that we hardware support makes the other computations slower. Since our objective is to speed up the computations, the only reason why we can tolerate a little slow-down of all other operations is when this particular operation becomes ***much faster*** than the one that we have used before.

Therefore, the following idea helps to decide which operations to support:

- If the direct hardware support makes an operation much faster, then it is probably desirable to directly support it in hardware.
- If, on the other hand, the direct hardware support, would not drastically speed u p this operation, then this operation is, probably, not the best candidate for hardware support.

Of course, this is simply a raw idea, and in the real applications, we have to actually compare the speed-ups and slow-downs to make a decision, but this idea gives a picture of how such a decision can be made.

4 At First Glance, From the Theoretical Viewpoint, Adding 1 Is Not That Different from General Addition, So There Seems To Be No Reason to Implement This Operation Separately

What we must do. From the viewpoint of this general idea, in order to decide whether adding 1 must be separately hardware supported or not, we must check whether the hardware support of adding 1 will really speed this operation up as compared to its implementation as particular case of the general integer addition.

In other words, we needs to compare the computation time of the following two operations:

- (general) addition of two integers
- adding 1 to an integer.

How can we estimate the computation time of each operation?

Operation time reformulated in computer terms. On the hardware level, each elementary hardware operation is an operation with bits. Therefore, the computation time of each operation can be estimated, crudely speaking, as proportional to the total number of *bit operations* that we have to perform.

The first seemingly natural choice: worst-case complexity. The number of bit operations depends on what exactly numbers we add. So, in order to compare two different algorithms that can be applied to different numbers, we must come up with some natural characteristics of these algorithms that will characterize their overall behavior.

In theoretical computer science (see, e.g., [1]), a typical measure of time complexity of an algorithm is its *worst-case* number of computational steps, i.e., the maximal number of steps that this algorithm needs to process the input of a given length n.

In the computer, usually, integers require a fixed number b of bits (e.g., in PC's, usually, 2 bytes = 16 bits are reserved for each normal size integer). So, we are interested in the worst-case bit complexity of

- adding two numbers of size h (i.e., integers with b bits in each of them), and
- adding 1 to a b-bit number.

Worst-case complexity of adding two integers. To add two b-bit integers, we must add their last **bits**, then add the previous bits (and maybe a carry), etc. If we count adding a *carry,* then we need 2 bit additions per each bit. Thus, out of b bits, in the worst case, we need 2 hit operations per bit except for the very last bit that requires only one operation. So, the worst-case bit complexity (= number of bit operations) of adding two numbers is $2(b-1)+1 = 2b-1$.

Worst-case complexity of adding 1 to an integer. The worst case of adding 1 is when we add 1 to a number 11...1 that consists of all 1 's. In this case, adding 1 will change all the bits in the number (to 100...0). Changing each bit requires at least one bit operation, so, in the worst case, adding 1 requires $b-1$ bit operations.

From the viewpoint of the worst-case complexity, the gain is minimal. Comparing these two results, we conclude that from the viewpoint, of the worst-case complexity, adding 1 takes, at best, half a time of the general addition.

This is faster, but, since adding each hardware operation increases the running time of other operations, this increase does not seem to necessarily justify the separate implementation of adding 1.

5 If We Consider a (More Meaningful) Average Time Instead of the Worst-Case Time, Then We Get the Drastic Difference Between Adding 1 and General Addition

Average-case complexity is more meaningful for our problem than the worst-case complexity. Our main objective is to make the computers faster. From the viewpoint of this objective, bad worst-case time is tolerable is the corresponding situations are rare. What we are really interested in is the *average* bit complexity, averaged over all possible numbers.

We will now show that, in contrast to the worst-case complexity, the average complexity does explain why adding 1 is sometimes implemented as a separate operation.

Average-case complexity of adding two integers. It is easy to show that an average-case bit complexity of adding two integers is still a linear function of b.

Average-case complexity of adding 1 to an integer: deducing a formula. The natural algorithm of adding 1 is as follows: We start at the lower digits. Whenever we encounter 1, we change it to 0. When we encounter O, we change it to 1 and stop. For this algorithm, the total number of bit operations is equal to the number of bits that are changed.

For example, if the number ends in O (e.g., 11 10), we replace this 0 by 1 and stop. which take exactly 1 bit operation. If the number end in 01 (e.g., 11 01), we replace two last bits and stop, making it two bit operations.

Let us calculate the average number of bit operations:

- For numbers ending in a single O, we need 1 bit operation. Such numbers constitute 1/2 = 2| of all numbers of size b.
- For numbers ending with a O and one 1 (i.e.. with 01), we need 2 bit operations. Such numbers constitute 1/4 = 2^{-2} of all numbers of size b.
- For numbers ending with a 0 and two 1's (i.e., with 011), we need 3 hit operations. Such numbers constitute 1/8 = 2^{-3} of all numbers of size b.

...

- For numbers ending with a O and $(k-1)$ 1's (i.e., with 01 . . . 1), we need 2 bit operation. Such numbers constitute 2^{-k} of all numbers of size b.

...

- For numbers ending in a single O and $(b-1)$ $1's$ (i.e., for a number 01 . . . 1), we need b bit operation. There is only one such number, and it therefore constitutes 2~ of all numbers of size b.
- The only remaining number is 1 ...1 (all 1 's), for which we also need b bit operations. This number also constitutes 2^{-b} of all numbers of size b.

The resulting average-time complexity $c(b)$ of adding 1 to a b-bit number is equal to

$$c(b) = 1\cdot 2^{-1} + 2\cdot 2^{-2} + 3\cdot 2^{-3} + \ldots + b\cdot 2^{-b} + b\cdot 2^{-b}.$$

Average-case complexity of adding 1 to an integer: simplifying a formula. For large b, it is natural to estimate this sum by considering the infinite sum

$$c(\infty) = 1\cdot 2^{-1} + 2\cdot 2^{-2} + 3\cdot 2^{-3} + \ldots + b\cdot 2^{-b} + (b+1)\cdot 2^{-(b+1)} + \ldots$$

Since

$$2^{-b} = 2^{-(b+1)} + 2^{-(b+2)} + \ldots$$

and $b \le b+1$, $b \le b+2, \ldots$ we can conclude that

$$b\cdot 2^{-b} = b\cdot(2^{-(b+1)} + 2^{-(b+2)} + \ldots) = b\cdot 2^{-(b+1)} + b\cdot 2^{-(b+2)} + \ldots \le (b+1)\cdot 2^{-(b+1)} + (b+2)\cdot 2^{-(b+2)} + \ldots.$$

and that, therefore, $c(b) \le c(\infty)$.

The value

$$c(m) = \sum_{k=1}^{\infty} k\cdot 2^{-k} \tag{1}$$

can be easily computed if we multiply both sides of this equality by 2 and take into consideration that $2\cdot 2^{-k} = 2^{-(k-1)}$:

$$2\cdot c(\infty) = \sum_{k=1}^{\infty} k\cdot 2\cdot 2^{-k} = \sum_{k=1}^{\infty} k\cdot 2^{-(k-1)}.$$

Introducing a new variable $j = k - 1$, we conclude that

$$2c(\infty) = \sum_{j=0}^{\infty}(j+1)\cdot 2^{-j} = \sum_{j=0}^{\infty} j\cdot 2^{-j} + \sum_{j=0}^{\infty} 2^{-j} \tag{2}$$

The first sum in the right-hand side of the equation (2) differs from the expression (1) only by the term corresponding to $j = 0$, which is equal to 0; thus, the first sum is equal to $c(\infty)$. The second sum $1 + 2^{-1} + 2^{-2} + \ldots$ is a geometric progression, its sum is 2. Hence, from (2), we can conclude that $2c(\infty) = c(\infty) + 2$ and $c(\infty) = 2$.

Thus, $c(\infty) = 2$, and $c(b) \leq 2$.

Average-case complexity of adding 1 to an integer: the result. As a result, we conclude that on average, adding 1 takes at most 2 bit operations, while adding two generic integers takes at least b bit operations. For 16-bit integers, it means a 8 times speed up. For double-size integers (with 32 bits) it means 16 times speed up, etc.

Conclusion. On average, ***adding 1*** is ***much*** *faster* than ***the general*** addition *of two* ***integers. This*** drastic speed- up explains why *the operation of* adding *1* is often separately hardware ***supported, and why this*** operation is often separately described in high-level ***programming languages like*** C++.

Acknowledgments. This work was supported by the NASA Pan American Center for Environmental and Earth Studies (PACES). The author is thankful to Ann Gates, Vladik Kreinovich, Luc Longpré, and Scott Starks for their help.

References

[1] C. H. Papadimitriou, ***Computational Complexity,*** Addison Wesley, San Diego, 1994.

Page intentionally left blank

URC97055

Fast Fuzzy Arithmetic operations

Michael Hampton[1] and Olga Kosheleva[2]

Departments of [1]Computer Science and
[2]Electrical and Computer Engineering
The University of Texas at El Paso
El Paso, TX 79968, USA
emails [1]mhampton@cs.utep.edu
[2]olga@ece.utep.edu

Abstract

In engineering applications of fuzzy logic, the main goal is not to simulate the way the experts really think, but to come up with a good engineering solution that would (ideally) be better than the expert's control, In such applications, it makes perfect sense to restrict ourselves to simplified approximate expressions for membership functions. If we need to perform arithmetic operations with the resulting fuzzy numbers, then we can use simple and fast algorithms that are known for operations with simple membership functions.

In other applications. especially the ones that are related to humanities, simulating experts is one of the main goals. In such applications, we must use membership functions that capture every nuance of the expert's opinion; these functions are therefore complicated, and fuzzy arithmetic operations with the corresponding fuzzy numbers become a computational problem.

In this paper, we design a new algorithm for performing such operations. This algorithm is applicable in the case when negative logarithms — $\log(\mu(x))$ of membership functions $\mu(x)$ are convex, and reduces computation time from $O(n^2)$ to $O(n \log(n))$ (where n is the number of points x at which we know the membership functions $\mu(x)$).

1 Formulation of the Problem

Depending on the goal, applications of fuzzy logic can be naturally divided into two classes:

- Engineering applications like *fuzzy control* in which fuzzy logic is used as a tool for achieving a certain goal: a better (smoother and safer) control of a car, a better heating, etc. In such applications, the expert's knowledge described by fuzzy rules is *used* not to *simulate* the way experts solve these *problems, but to design better control strategies,*

- Applications to humanities (psychology, linguistics, etc.) in which fuzzy logic is used to *describe* and simulate the human behavior, the human decision-making processes, etc., and thus *predict* the way humans will react in different situations.

In both types of applications, we have to deal with *fuzzy numbers r*, i.e., quantities whose values we do not know precisely, and instead, we only have expert (fuzzy) knowledge about these values. This knowledge is usually described in terms of *membership functions* $\mu_r(x)$ that assign to every real number x the expert's degree of belief $\mu_r(x) \in [0, 1]$ that the actual (unknown) value of the quantity r is equal to x.

The formalism (membership functions) is the same, but, depending on the application, we treat these membership functions differently:

- In engineering applications, we do not need to describe the *exact* opinion of the experts, because we are going to *improve* this description (by some fine-tuning) anyway. Therefore, it is quite sufficient to use membership functions that *approximately* describe expert's opinions. To simplify computations, usually, the simplest approximations are used, most often triangular or trapezoid membership functions (see, e.g., [2]).

- In humanities applications, if we use oversimplified approximations to membership functions, we will end up having very crude models of human behavior For such applications, we, therefore, need accurate descriptions of membership functions. and these descriptions can be very complicated.

1.1 Fuzzy Data Processing and Fuzzy Arithmetic Operations: If We Must Use Precise Membership Functions, We Have a Computational Problem

Fuzzy data processing. We want to use the expert (fuzzy) knowledge about the values $r_1, \ldots, r_n$ of some quantities to *predict* the value of some quantity r that is related to r_i. In this paper, we will consider the simplest case when "related" means that we know the exact form of the dependency $r = f(r_1, \ldots, r_n)$ between r_i and r, and the only uncertainty in r is caused by the uncertainty in the values of r_i.

For example, when we formalize the expert's opinion about possible candidates for a position, we may know that this opinion depends on the values of n characteristics r_i of the candidate, we have expert (fuzzy) knowledge about the values of r_i, and we know that the final opinion depends on the total evaluation $r = w_1 \cdot r_1 + . + w_n r_n$ with known weights w_i.

In such situations, we must transform the fuzzy knowledge about the values r_i into a fuzzy knowledge about $r = f(r_1, \ldots, r_n)$. This transformation is called *fuzzy data processing*.

Fuzzy arithmetic operations. In the computers, usually, only elementary arithmetic operations ($+$, $-$, $\cdot$, $/$) are hardware supported. Therefore. every data processing algorithm written in a high-level programming language is *parsed*, i.e., represented as a sequence of elementary arithmetic operations. For example, computing an expression $x_1(x_2 + x_3)$ is decomposed into two steps: Computing $x_2 + x_3$ and multiplying the result by x_1.

In view of this decomposition, in order to implement. an arbitrary data processing algorithm with fuzzy inputs, it is sufficient to be able to apply *elementary arithmetic operations* $o = +, -, \cdot$, to fuzzy numbers. The formulas for these operations come from the *extension principle (see, e.g.*, [3]): In particular, if we use an algebraic product $a \cdot b$ as a fuzzy analogue of &, we arrive at the following formula for $t = r \circ s$:

$$\mu_t(x) = \sup_{y,z\ y \circ z = x} (\mu_r(y)\ \mu_s(z)). \tag{1}$$

In particular, for $\circ = +$, we have

$$\mu_t(x) = \sup_y (\mu_r(y) \cdot \mu_s(x-y)). \tag{2}$$

For simple membership functions, fuzzy arithmetic operations are computationally easy. For example, if we use Gaussian membership functions

$$\mu_r(x) = \exp((x - a_r)^2/(\sigma_r)^2),$$

$$\mu_s(x) = \exp((x - a_s)^2/(\sigma_s)^2),$$

then *(2) leads to* a Gaussian *membership function for t: $\mu_t(x) = \exp((x - a_t)^2/(\sigma_t)^2)$ with*

$$a_t - \frac{a_r(\sigma_r)^{-2} + a_s(\sigma_s)^{-2}}{(\sigma_r)^{-2} + (\sigma_s)^{-2}}$$

and $(\sigma_t)^{-2} = (\sigma_r)^{-2} + (\sigma_s)^{-2}$ [3, 5]. These are computationally very simple formulas to implement.

There are simple formulas for several other cases (see, e.g., [3] and references therein).

For complicated membership functions, fuzzy arithmetic operations are computationally complicated. When we cannot use approximating simple expressions, then we cannot use simplified formulas that stem from the use of these expressions, and therefore, we have to use the formula (2). This formula is straightforward, so, we can simply use it to compute $\mu_t(x)$. To find out how long it would take to compute $\mu_t(x)$, let us estimate the number of computational steps that are required to compute $\mu_t(x)$.

Of course, in reality, we can only know the values of $\mu_r(x)$ and $\mu_s(x)$ for finitely many values x. Let us denote the total number of such values by n. In this case, it is reasonable to compute only n values of $\mu_t(x)$. For each of these n values, according to the formula (2), we must find the largest of n products. Computing each product takes 1 elementary computational step, computing the largest of n numbers requires that we do $n - 1$ comparisons. So, the total number of computation steps that needs to be done to compute one value of $\mu_t(x)$ is $2n - 1 = O(n)$.

If we have n parallel processors at our disposal. then we can use each processor Lo compute its **own** value of $\mu_t(x)$ and thus, compute *all* these values in linear time.

In many real-life situations, however, we only have one computer. In such situations, to compute *all n* **values** of the desired membership function $\mu_t(x)$, we need $O(n^2)$ computational steps.

The more accurately we wish to represent the expert's opinion, the larger n we need to take. For large $n, O(n^2)$ is too long. Can *we perform fuzzy arithmetic operations faster?*

In [4], an *approximate* algorithm is given that performs arithmetic operations with fuzzy numbers in time $O(n\log(n))$.

1.2 What We Are Planning to Do

In this paper, we design a new fast algorithm that computes the precise value of the resulting membership functions in $O(n\log(n))$ time.

This algorithm is applicable when the negative logarithms $-\log(\mu(x))$ of the membership functions $\mu(x)$ are convex. This class of membership functions includes many important classes such as Gaussian membership functions.

2 Fast Addition of Fuzzy Numbers

2.1 Main Idea

Let us describe, step-by-step, how we can simplify the problem of computing the sum of two fuzzy numbers.

First simplification: reformulation in discrete terms. We only know the membership functions p,(x) and $\mu_s(x)$ in finitely many points. and usually, these points are of the type $x_i = i$ Ax. **In** this case, the formula (2) takes the following form:

$$t_i = \max_j (r_j\, s_{i-j}), \tag{3}$$

where we denoted $t_i = \mu_t(i\ \Delta x), r_i = \mu_r(i\,.\,\Delta x)$, *and* $s_i = \mu_s(i\,\Delta x)$.

Further simplification: reducing multiplication to addition. The formula (3) can be simplified even further if we recall that the equality $t = r \cdot s$ is equivalent to 'T = R + S, where $T = -\ln(t)$, R = $-\ln(r)$, and S = – in(s). In view of this equivalence, and taking into consideration the fact that – in(z) is a strictly decreasing function, we can reformulate the formula (3) as follows:

$$T_i = \min_j (R_j + S_{i-j}), \tag{4}$$

where we denoted $T_i = -\ln(t_i), R_i = -\ln(r_i)$, and $S_i = -\ln(s_i)$. We will describe how, given the two sequences R_i and S_i, we will be able to compute the elements T_i fast. Then, if we know the values $r_i = \mu_r(i\cdot\Delta x)$ and $s_i = \mu_s(i\,\Delta x)$, we will be able to compute the values R_i and S_i, compute $T_i = -\ln(t_i)$, and then reconstruct the desired values $t_i = \mu_t(i\,\Delta x)$ **as** $t_i = \exp(-T_i)$.

How to compute the formula (4)?

Final simplification: a local criterion for the maximum. For a given i, when does the sum $\sum_j = R_j + S_{i-j}$ attains its minimum'? If it does attain the minimum for some j, this means that the value of this sum for this particular j is not larger than the values of this sum for $j-1$ and for $j+1$: $\sum_j \le \sum_{j-1}$ and $\sum_j \le \sum_{j+1}$. If we denote $D_j = \sum_j - \sum_{j-1}$, then these two inequalities take the form

$$D_j \le 0; \quad D_{j+1} \ge 0. \tag{5}$$

We can use binary search to find the desired j. Since the function $-\ln(\mu_r(x))$ is convex, the sequence R_j is also convex, and therefore, the differences $R_j - R_{j-1}$ are monotonically increasing with j. Similarly, the differences $S_{i-(j-1)} - S_{i-j}$ are strictly *decreasing* with j. Therefore. the difference $D_j = (R_j - R_{j-1}) - (S_{i-j} - S_{i-(j-1)})$ is increasing with j.

Hence, we can find the desired value j that satisfies the condition (5) by **using** *binary search:* This will be an iterative process on which, on each step, we will have lower and upper bounds for the desired value j. We start with the lower and upper bounds that encompass all possible values of j. Then, on each iteration. we:

- take a midpoint $m = (\texttt{lower} + \texttt{upper})\,\texttt{div}\,2$ between the current lower and the upper bounds;
- compute D_m for this midpoint m, and
- compare the resulting value D_m with 0.

Depending on the result of this comparison, we do the following:

- If $D_m = 0$, then, due to the monotonicity of the sequence D_m, we **have** $D_{m+1} \geq D_m =$ **0, i.e.,** $D_{m+1} \geq$ **0.** Hence, this m satisfies the condition (5). Using monotonicity of D_j, one can easily show that in this case,
  - either m is the only value for which (5) in true (in which case, it is the only possible minimum of $\sum_m$),
  - or $D_j = 0$ not only for $j = m$, but also for several values of j that are neighboring to m, in which case, there are several minima with exactly the same value of $\sum_j$.

  In both cases, the value of $\sum_m$ for the midpoint m *is* the desired minimum.
- If $D_m > 0$, this means, due to monotonicity of the sequence D_j, that $j < m$. In this case, we can take m *as* the new value of the variable upper.
- Similarly, if $D_m < 0$, this means, due to monotonicity of the sequence D_j, that $m < j$. In this case, *we* can take m as the new value of the variable lower.

This algorithm takes $O(n$ **log(n)) steps. On each iteration of the binary** search, we reduce the size in half. In k iteration, we go down from n to $\leq n/2^k$ possible values. When $n/2^k \leq 1$, we are down to a single point, and thus, we have localized the desired j. The inequality $n/2^k \leq 1$ is achieved when $k \approx \log_2(n)$, so, we need $O(\log(n))$ points to find the desired j and thus, to compute the desired value of T_i for this particular i.

To compute the values of T_i for n different i's, we thus need $n \cdot O(\log(n)) = O(n \log(n))$ computational steps.

2.2 Resulting Algorithm

GIVEN: the values p,(r) and $\mu_s(x)$ for n equally spaced values $x_i =$ r' . Δx.

A*LGORITHM:*

- First, for each of n values x_i, we compute the values $R_i = -\ln(\mu_r(x_i))$ and $S_i = -\ln(\mu_s(x_i))$.
- For each i, we:
  - apply binary search to find the index j for which the non-decreasing sequence $D_j = (R_j - R_{j-1}) - (S_{i-j} - S_{i-(j-1)})$ passes from the non-positive to non-negative values;
    compute T_i as $R_j + S_{i-j}$ for this very j;
  - compute $\mu_t(x_i)$ as $\exp(-T_i)$.

3 Algorithms for Other Arithmetic Operations

3.1 Subtraction

To compute $t = r$ –s, we can represent it as $t = r + (-s)$. Since we know the membership function $\mu_s(x)$ for s, we can easily compute the membership function $\mu_{-s}(x)$ for –s as $\mu_{-s}(x) = \mu_s(-x)$. Then, we can apply the above algorithm to compute the desired membership function for $t = r - s = r + (-s)$.

3.2 Multiplication

If the quantities r and s both take only positive values, then, to compute r s, we can use the formula $r\ s = \exp(\ln(r) + \ln(s))$:

- From the membership functions for r and s, we can easily compute the membership functions for $\ln(r)$ and $\ln(s)$ as $\mu_{\ln(r)}(x) = \mu_r(\ln(x))$ and $\mu_{\ln(s)}(x) = \mu_s(\ln(x))$.
- Applying the algorithm presented above, we compute the membership function $\mu_{\ln(t)}$ for $\ln(t)$ = In(r)+ $\ln(s)$.
- Finally, from $\mu_{\ln(t)}$, we compute $\mu_t(y)$ as $\mu_t(y) = \mu_{\ln(t)}(\exp(y))$.

3.3 Division

Division $t = r/s$ *can he* expressed as $t = r(1/s)$. So, to divide two fuzzy numbers, we can use the following algorithm:

- First, we compute the membership function for $1/s$ as $\mu_{1/s}(x) = \mu_s(1/x)$.
- Then, we use the algorithm for multiplication to compute the membership function for

$$t = r \cdot (1/s) = r/s.$$

3.4 Computational Complexity

For all **these** operations, the major part is computing the sum of fuzzy numbers that takes $O(n\ \log(n))$ steps. Therefore, the computational complexity of computing the difference, product, or ratio of two fuzzy numbers is also $O(n \log(n))$.

4 What If A t-Norm (&−Operation) Is Different From Algebraic Product?

4.1 Fuzzy Arithmetic Operations: Case of a General t-Norm

For an arbitrary &—operation $f_\&(a, b)$, the extension principle for addition leads to the following formula:

$$\mu_t(x) = \sup_Y f_\&(\mu_r(y), \mu_s(x - y)). \tag{3}$$

4.2 Strictly Archimedean t-Norms and Reduction to the Case of Algebraic Product

Idea. [t is known (see, e.g., [3]), that if an &−operation satisfies some reasonable conditions, then it **can** represented i n the form

$$f_\&(a, b) = \psi^{-1}(\psi(a) \,.\, \psi(b)) \tag{4?}$$

for some strictly increasing function $\psi : [0,1] \rightarrow$ **[0, 1]** (&−operations **that satisfy these "reasonable"** conditions are called *strictly Archimedean*).

Since the function ψ is strictly increasing, the value $f_\&(\mu_r(y), \mu_s(x - y))$ is the largest iff the value $\psi(f_\&(\mu_r(y), \mu_s(x - y)))$ is the largest, so.

$$\psi(\mu_t(x)) = \sup_Y \psi(f_\&(\mu_r(y), \mu_s(x\ y)). \tag{5}$$

From (4), we conclude that $\psi(f_\&(\mu_r(y), \mu_s(x-y))) = \psi(\mu_r(y)) \cdot \psi(\mu_s(x-$ y)). Therefore, (7) can be rewritten as:

$$\psi(\mu_t(x))^{-} \sup_y \psi(\mu_r(y))\ \psi(\mu_s\,(x\ -\ y)). \tag{6}$$

If we denote $\nu_r(x) = \psi(\mu_r(x)), \nu_s(x) = \psi(\mu_r(x))$, and $\nu_t(s) = \psi(\mu_t(x))$, then this formula will take the form

$$\nu_t(x) = \sup_Y(\nu_r(\,y)\,\nu_s(x\ Y\,))\,! \tag{7}$$

which is exactly like the formula (2) that we already know how to compute fast. From $\nu_t(x) = \psi(\mu_t(x))$, we can compute $\mu_t(x)$ by applying an inverse function ψ^{-1}: $\mu_t(x) = \psi^{-1}(\nu_t(x))$.

So, to compute $\mu_t(x)$, we can apply the following algorithm:

Algorithm.

- For every x, compute $\nu_r(x) = \psi(\mu_r(x))$ and $\nu_s(x) = \psi(\mu_s(x))$. This takes $O(n)$ steps.
- Apply the algorithm (described in the previous section) to $\nu_r(x)$ and ν_s (x); this algorithm will take $O(n$ log(n)) computational steps and return $nu_t(x)$.
- Apply the inverse function ψ^{-1} to $\nu_t(x)$, resulting in $\mu_t(x) = \psi^{-1}(\nu_t(x))$. This is done value-by-value. so, for $O(n)$ values of x, it takes $O(n)$ steps.

Computational Complexity. The resulting algorithm requires

$$o(n) + O(n\log(n)) + O(n) = O(n\log(n))$$

computational steps.

4.3 Other Arithmetic Operations

For other arithmetic operations with fuzzy numbers (–, . . /), we have a similar reduction to the case of algebraic product that leads Lo similar $O(n$ log(n)) algorithms.

Acknowledgments. This work was supported by the Office of Naval Research Grant No. N00014-93-1-1343 and, partially, by the National Science Foundation Grant No. CDA 9522903, and by the NASA Pan American Center for Environmental and Earth Studies (PACES). Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the view of the funding agencies.

The authors are thankful to Ann Gates, Vladik Kreinovich, Luc Longpré, and Scott Starks for their encouragement.

References

[1] Th. H. Cormen, Ch. L. Leiserson, R,. L. Rivest, *Introduction to algorithms*, MIT Press. Cambridge, MA, 1990.

[2] K. Hirota and M. Sugeno. *industrial Applications of Fuzzy Technology in the World.* World Scientific, Singapore, 1996.

[3] G. Klir and B. Yuan, *Fuzzy sets and fuzzy logic: theory* and *applications*. Prentice Hall, Upper Saddle River, NJ, 1995.

[4] O. Kosheleva, S. D. Cabrera, G. A. Gibson. and M. Koshelev, "Fast Implementations of Fuzzy Arithmetic Operations Using Fast Fourier Transform (F FT)", *Proceedings of the 1996 IEEE International Conference on Fuzzy Systems. New* Orleans. September 8-11, 1996, Vol. 3, pp. 1958-1964.

[5] V. Kreinovich, C. Quintana, and L. Reznik. Gaussian membership functions arc most adequate in representing uncertainty in measurements. *Proceedings of NA FIPS'92: North American Fuzzy Information* Processing Society *Conference, Puerto Vallarta*. Mexico, *December 15-17, 1992*, NASA Johnson Space Center, Houston, TX, 1992. pp. 618-625.

URC97056

Piecewise Linear Membership Function Generator - Divider Approach

Ron Hart, Gene Martinez, Bo Yuan, Djuro Zrilic, Jaime Ramirez[1]

NASA's Center for Autonomous Control Engineering

Department of Engineering

New Mexico Highlands University

Las Vegas, New Mexico 87701

Abstract- In this paper a simple, inexpensive, membership-function circuit for fuzzy controllers is presented. The proposed circuit may be used to generate a general trapezoidal membership function. The slope and horizontal shift are fully programmable parameters.

1. Introduction

There exist numerous examples of industry proven fuzzy-logic solutions in the domain of control, expert systems, and pattern recognition. A typical Mamdani fuzzy controller consists of a fuzzifier block, an inference-engine block, a rulebase block, and a defuzzification block [1]. The hardware realization of the membership-function generator circuit is the subject of this paper.

The proposed circuit is characterized by a generalized trapezoidal membership function defined in terms of horizontal shifting parameters. The authors realized a parametric class of membership functions based on piecewise linear functions. In this paper, they present a basic mathematical background of membership functions, second, simulation results for the current design, and the results for hardware realization.

2. Proposed Algorithm

We propose to realize the following function.

$$F(v) = \begin{cases} 0 & if \quad v < V_a \\ \dfrac{v - V_a}{V_b - V_a} & if \quad V_a \le v < V_b \\ 1 & if \quad V_b \le v \le V_c \\ \dfrac{v - V_d}{V_c - V_d} & if \quad V_c < v \le V_d \\ 0 & if \quad V_d < v \end{cases} \tag{1}$$

where Va, Vb, Vc, and Vd are real numbers determining the shape of the trapezoid illustrated in Figure 1.

[1] Klipsch School of Engineering, New Mexico State University, Las Cruces, NM, 88003

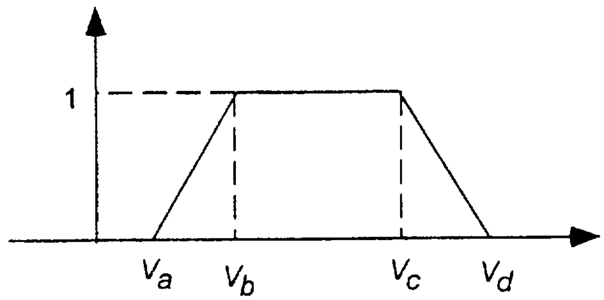


Figure1 A Trapezoidal Membership Function

To realize expression (1), two analog multiplexers, two summers, and a divider are needed. A block diagram of the proposed circuit is shown in Figure 2. Threshold voltages are multiplexed then combined with the input signal according to equation (1). To realize the required voltage ratio, a divider circuit consisting of an AD534JD is used.

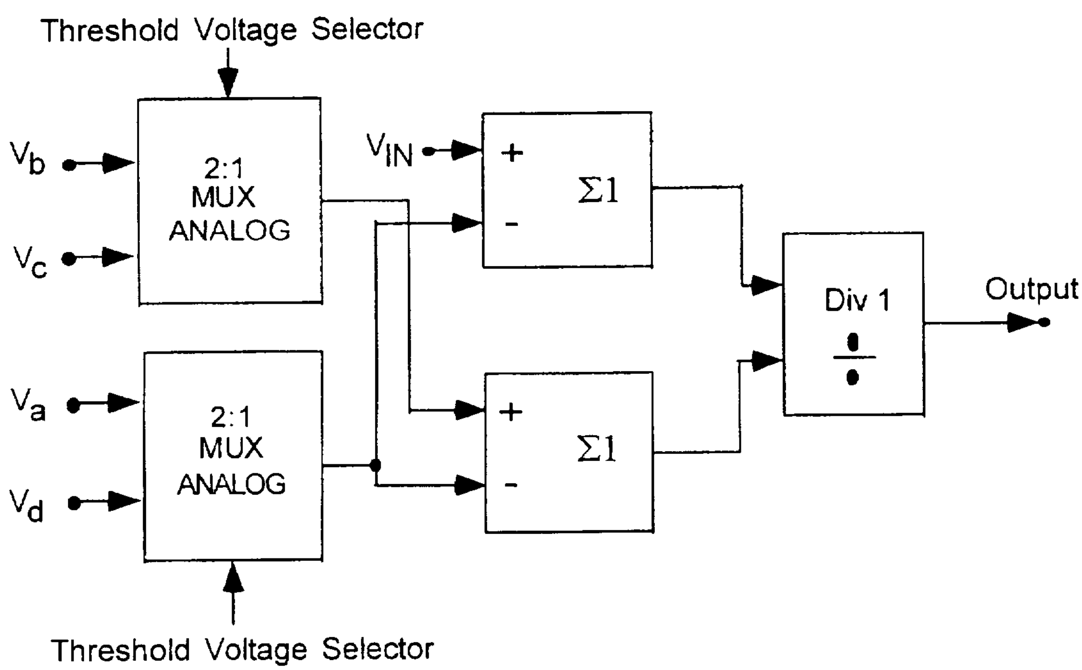


Figure 2 Block diagram of the proposed membership function circuit.

To validate this approach prior to hardware realization, computer simulations were performed. These are discussed in the next section.

3. Simulation Results

In our simulations, the user friendly software package TESLA is used [2]. This is a modular software system consisting of analog and digital components needed to perform this simulation. The simulation block diagram with numbered nodes is shown in Figure 3.

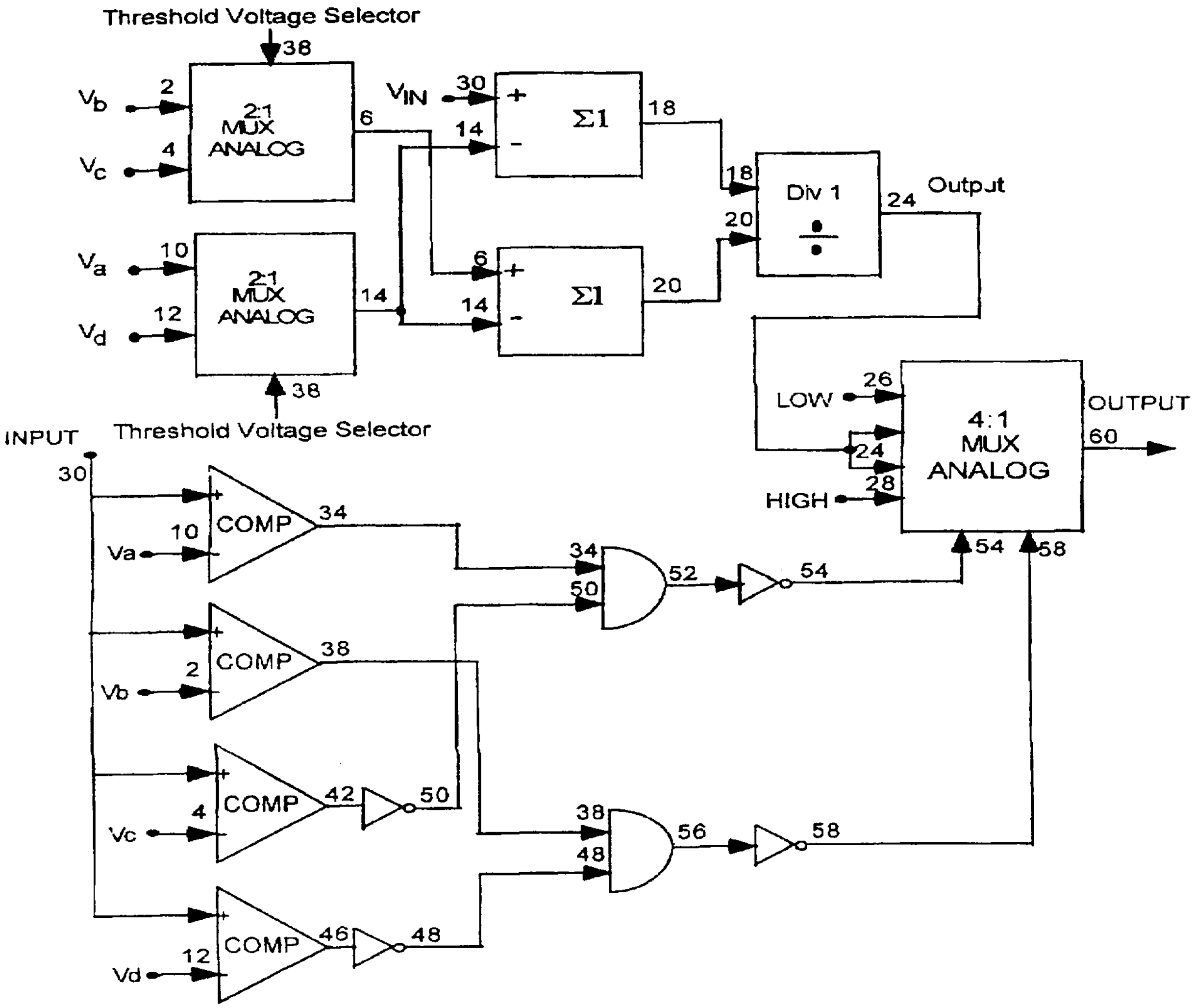


Figure 3 Simulation Block Diagram

The numbers found in the net-list in Table 1 correspond to the nodal connections in the block diagram.

X 0 0 30 0 FCNGEN FCN=3 F=.05 V=5
Va 10 PWR V=set(Va)=
Vb 2 PWR V=SET(Vb)=
Vc 4 PWR V=SET(Vc)=
Vd 12 PWR V=SET(Vd)=
MUX1 38 2 4 6 MUX
MUX2 38 10 12 14 MUX

SUM1 30 14 18 SUM G1=1 G2=-1
SUM2 6 14 20 SUM G1=1 G2=-1
DIV1 18 20 24 DIV
MUX4 54 58 26 24 24 28 60 MUX4
LOW 26 PWR V=0
HIGH 28 PWR V=1
COMP1 30 10 34 COMP
COMP2 30 2 38 COMP
COMP3 30 4 42 COMP
COMP4 30 12 46 COMP
NAND1 34 50 52 NAND
INV1 52 54 INV
NAND2 38 48 56 NAND
INV2 56 58 INV
INV3 42 50 INV
INV4 46 48 INV

Table 1: Simulation net list

The input signal was applied interactively along the horiontal voltage axis. For one set of simulations, horizontal parameters Va, Vb, Vc, and Vd are kept constant. Figure 4 presents the result of a point to point simulation for threshold voltages Va = 1V, Vb = 2V, Vc = 3V, and Vd = 5V.

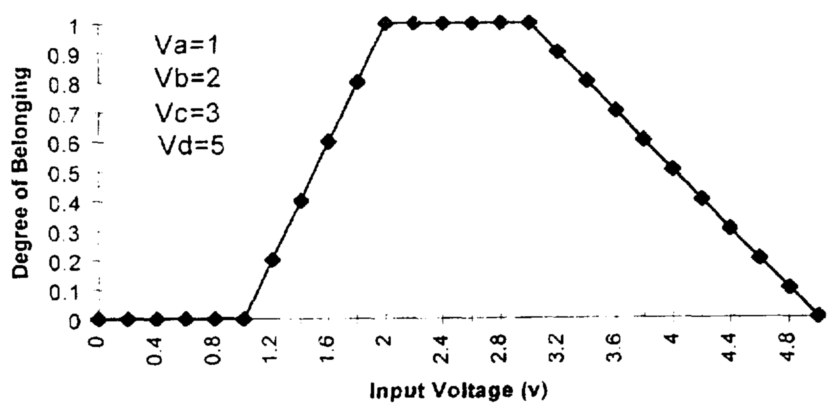


Figure 4 Simulated Trapezoidal Membership Function

The trapezoidal membership function can be considered as a general shape. For a triangular function, for example, Vb = Vc. Except, for horizontal shifting, the slope of each straight line can be changed by altering values of the threshold voltages as well. Figure 5 presents simulation results for some special trapezoidal shapes.

4. Realization Results

The block diagram of Figure 6 is a representation of the hardware realization of the simulated circuit shown in Figure 3. To verify our idea, inexpensive off shelf components were used to implement the circuit. As comparators we used UA741CP amplifiers. Analog Multiplexers are CD4066B bilateral switches. The divider circuit is the AD534JD. The threshold voltages Va, Vb, Vc, and Vd are adjusted to realize any symmetrical or asymmetrical shape. In Figure 7, two oscilloscope diagrams for different values of threshold voltages Va, Vb, Vc, and Vd are illustrated.

5. Conclusion

In many applications, real time fuzzy processors are needed. This paper presented an attempt to realize a simple inexpensive membership-function circuit. To verify our idea we performed simulations first, then realized the membership function generator using inexpensive off shelf components.

6. References

[1] Mamdani, E. E. [1974], "Applications of fuzzy algorithms for control of simple dynamic plant." Proc. IEE 121 (22), pp. 1585 - 1588.

[2] Tesoft, Inc. 205 Crossing Creek Ct, Roswell, GA 30076

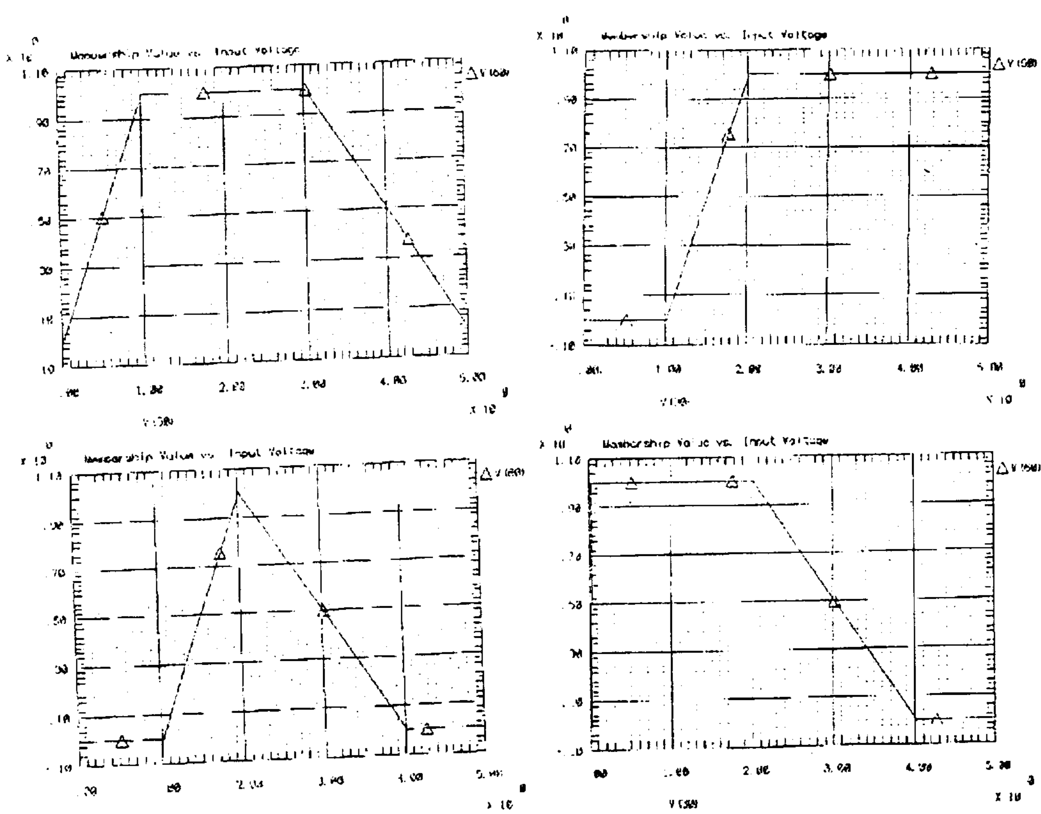


Figure 5 Plots of Simulation

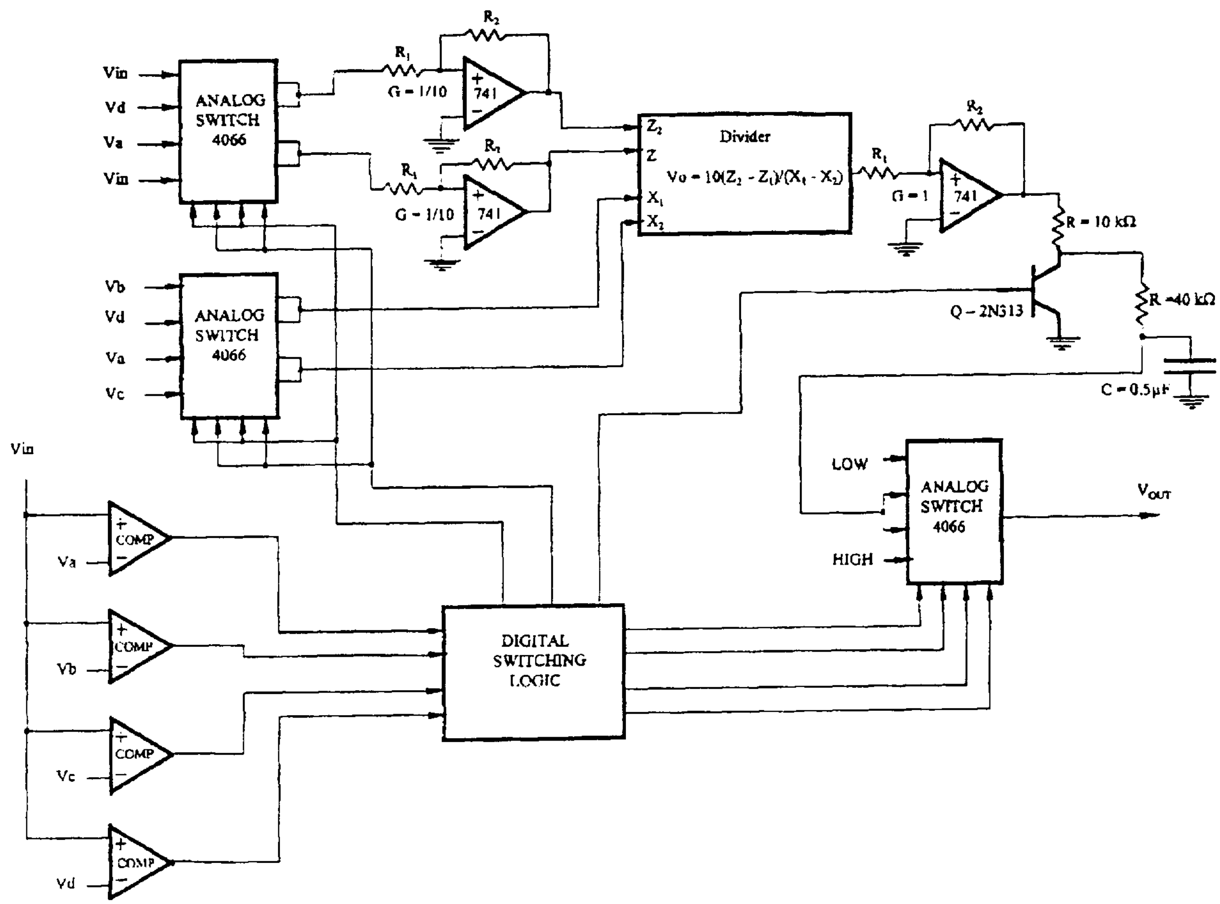


Figure 6 Circuit Block Diagram

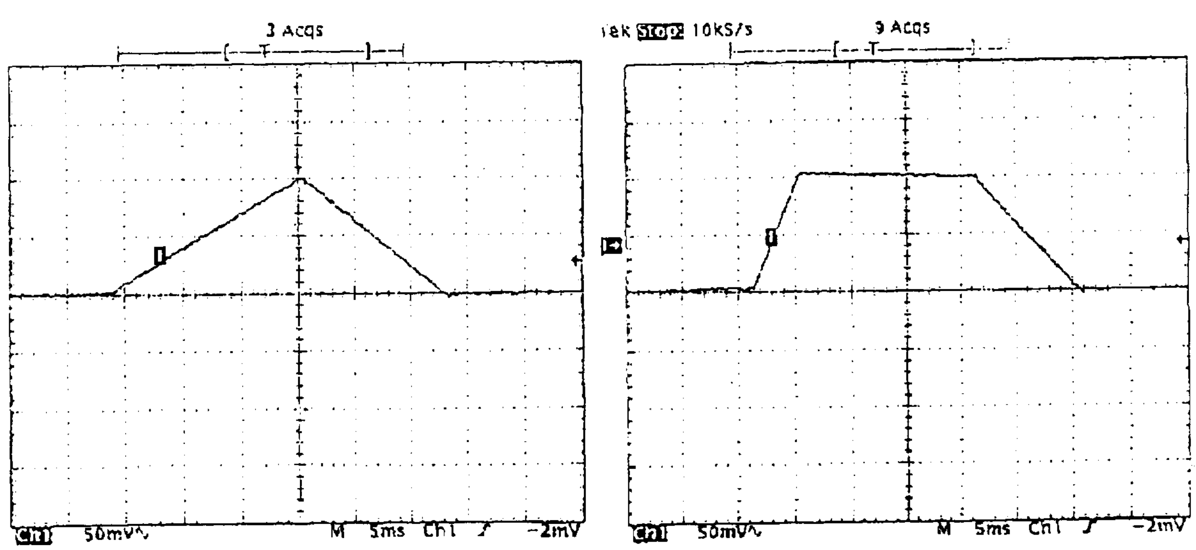


Figure 7 Realization Diagrams of Membership Function

URC97057

Preliminary Studies on Silicon Carbide Bulk Acoustic Wave devices

H.S. Henry, I. Minus, G.L. Harris, M.G. Spencer
Materials Science Research Center Of Excellence
Howard University

Introduction

In 1880, Pierre and Jacques discovered that applying pressure to a piece of quartz created an electrical potential between the deformed surfaces, and that the application of a voltage across the material caused physical displacements [1]. This effect is the piezoelectric effect and occurs only in ionic crystalline solids lacking a center of inversion symmetry. This property of quartz has been used to make devices such as pressure sensors, acoustic transducers, oscillators and microbalances.

Quartz bulk acoustic wave (BAW) devices are typically used for stable oscillator circuits and for quartz crystal microbalances (QCMs). These devices have overlapping electrodes on opposite sides of the quartz and operate in a thickness shear mode (TSM) meaning that the physical displacement is in a direction perpendicular to the electric field between the electrodes. The operating frequency of TSM devices is directly proportional to the acoustic velocity and inversely proportional the thickness of the material between the electrodes [2]. Thus the maximum operating frequency and device sensitivity is limited by a material's acoustic properties and the device thickness.

Silicon carbide (SiC) has an acoustic velocity almost one order of magnitude higher than that of quartz and is also a very sturdy material. This implies that it is possible to have devices operating at a much higher frequency than quartz devices for the same material thickness. It is with this in mind that we report on some preliminary studies on 3C and 4H polytypes of SiC for BAW devices. Table 1 shows the acoustic velocity for Quartz and SiC. The devices were characterized using a TTL-based oscillator circuit to observe the oscillator frequency and stability and the atomic force microscope (AFM) used to observe surface displacement.

Experimental

Devices were fabricated using single and polycrystalline 3C-SiC and Vanadium-doped 4H-SiC. The BAW devices tested were prepared by first doing a standard cleaning procedure. After cleaning, photoresist was spun on one side of the wafer, pre-baked and then exposed to UV light through a mask with the contact pattern. The resist was then developed, and a contact consisting of 200A Cr and 2000A Al was deposited using an electron beam (E-beam) evaporation system at a pressure of about 10^{-7} Torr. After a lift-off process to remove the excess metal, the same procedure was used to deposit the contacts on the opposite side of the wafer. Care was taken to make sure that the overlap of the contacts were fully matched. The overlapping area of the contacts were circles with a diameter of about 3.3mm. The finished sample were then either mounted in HC48/U crystal holders or prepared for use in the AFM.

The TTL-based oscillator used is based on a circuit given by Brukenstein and Shay[3]. The circuit was soldered in place on a printed circuit (PC) board with the devices placed as close together as possible. The PC board was then mounted in a metal enclosure to reduce stray effects. Figure 1 shows the two circuits used. The accuracy of both circuits were checked by using several commercial crystals with frequencies over the range 3.6-32MHz. Measurements show the circuit in figure 1 b to be accurate over the range 3-21 MHz while the circuit in figure 1a was accurate only above 20MHz. The upper limit on circuit could not be verified due to the lack of fundamental mode crystals above 32MHz. The frequency of operation was checked using both a Tektronix 2235 oscilloscope and a Tektronix CFG280 function generator with frequency counter.

Results and Discussion

It was anticipated that the poly SiC would be a good candidate for BAW devices. However, the results show this material to be not very good. The DC resistance for these devices were in the 100KΩ range compared to commercial systoles at greater than $10^{14}\Omega$. The oscillation frequencies from these devices were in the 10-20KHz range and the signal was quite noisy and unstable. The poly SiC devices were not measured on the AFM since they did not exhibit good oscillator characteristics. The material was about 500µm thick.

The single crystal 3C-SiC samples measured were previously fabricated by Griffin [4] and were already mounted in HC/6U holders. There were approximately 10µm thick. This made it impossible to test them on the AFM. These devices worked in the 18-45MHz range with most exhibiting stable oscillations. However, there is some concern with the results obtained from these devices. When measured in each of the circuits in figure 2, different frequencies were observed, many time both being quite stable. There is a possibility of each circuit exciting a different oscillation mode.

The implanted 4H-SiC sample performed much better than the other materials. Although the DC resistance could not be accurate] y measured, it lies in the range of 16- 100MΩ. When measured on a Kiethley electrometer, a measurement of 16MΩ was observed when set for autoranging. When the range was set manually, we measured resistances as high as 100MΩ. Doing the measurement on a HP3468A multimeter caused it to over-range, indicating the device had a resistance greater than 20MΩ.

The 4H sample gave an oscillation frequency of 46.9MHz which is much higher than predicted by the formula

$$f = \frac{V}{2d} \qquad (1)$$

where V is the bulk acoustic velocity and d is the material thickness. Based on (1), a frequency of 11.75MHz was anticipated. A possible explanation for this discrepancy is that the device was operating in an overtone mode and not in the fundamental mode.

Figure 2 shows the results from the AFM measurements. The lower trace shows the surface when no voltage is applied while the upper trace shows the surface after the application of a 10 volt potential. It is clearly seen that the application of a voltage across the device caused some surface deformation. The AFM will measure only vertical displacements, so it could not be verified whether or not the device was operating in the thickness shear mode as quartz does.

Conclusion

It has been demonstrated that SiC may be a useful material for acoustic wave devices. However, more work needs to be done in characterizing the devices and studying the acoustic properties.

References

1. C. Lu and A.W. Czanderna, Eds., Applications of Peizoelectric Quartz Crystal Microbalances, Elsever, Amsterdam, 1984.

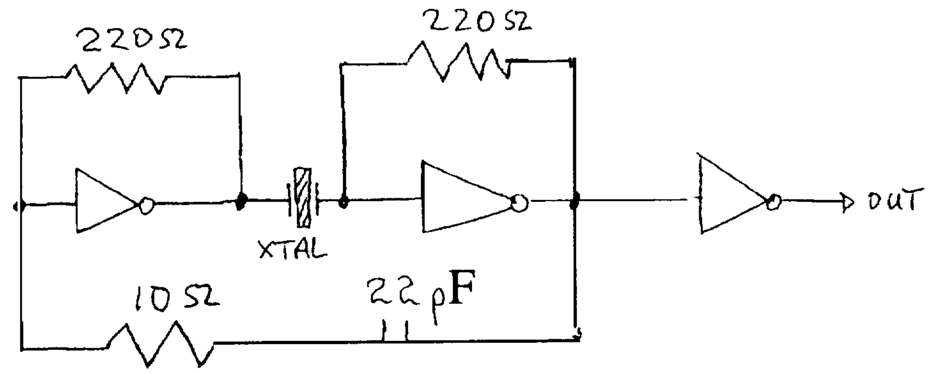
2. G. Sauerbrey, Z. Phys., 155, 206-222 (1 959)

3. S. Bruckenstein and M. Shay, Electrochimica Acts, 30, 1295-1300 (1985).

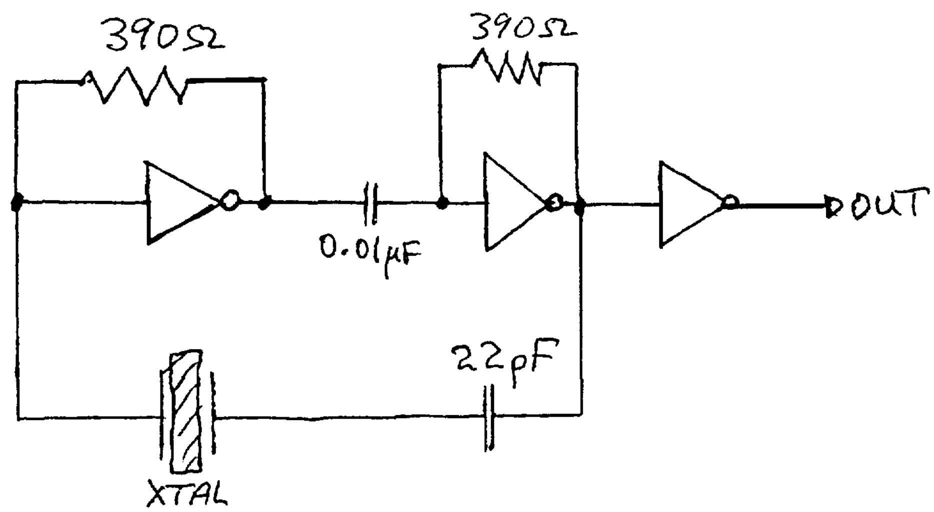
4. Ruben L. Griffin, Silicon Carbide Crystal Microbalance, Master's Thesis, Howard University, May 1995.

Table 1 Bulk acoustic velocity for Quartz and SiC

Quartz (km/s)	Silicon Carbide (km/s)
1,67 (a)	13.73 (4H-SiC)
	13.26 (6H-SiC)
	12.60 (3 C-SiC)



(a)



(b)

Figure 1 **TTL-based** oscillator circuits

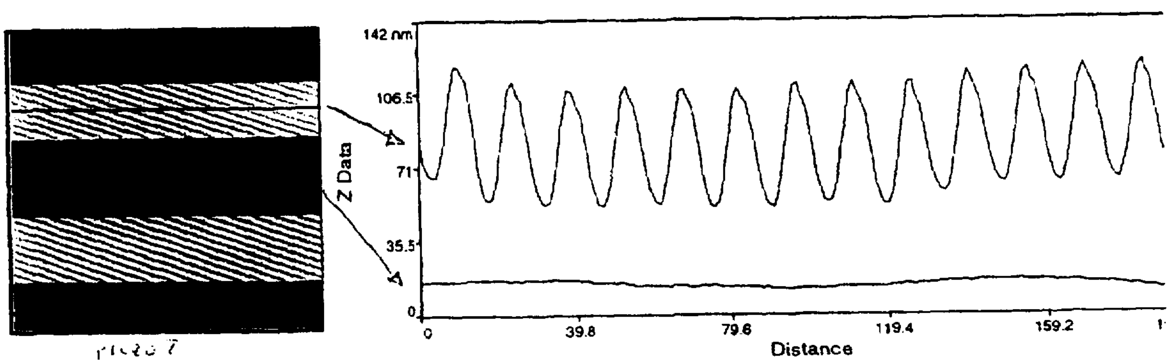


Figure 2 AFM results for Vanadium-doped the 4H-SiC sample.

Page intentionally left blank

URC97058

Numerical Simulation and Experimental Constraints of Bubble Growth in Rhyolite Glasses

S58/34

Donald M. Hooper and Glen S. Mattioli
Department of Geology
University of Puerto Rico at Mayagüez

Introduction

A quantitative description of bubble nucleation and growth rates is critical to our understanding various eruptive mechanisms and degassing processes. Volatile exsolution from a magmatic melt ultimately affects the eruptive style and degree of explosiveness. Bubble growth in magmas is primarily controlled by two processes: the diffusion of gas out of a supersaturated melt and the decompression of the gas inside the bubble as it rises and hydrostatic pressure decreases (Sparks, 1978). Bubbles may also grow by coalescence. An increase in the bubble radius is opposed by surface tension and resistance to growth due to the viscosity and inertia of the surrounding liquid. Viscosity, diffusion rate, surface tension, temperature, and dissolved volatile content are several important parameters which constrain both diffusional and decompressional growth. Here we report on a combined approach involving numerical simulations and new experimental observations to evaluate the existing theoretical framework.

Background

Diffusion is the process by which matter is transported from one part of a system to another down the chemical potential gradient as a result of random molecular motions (Crank, 1975). In most cases, this results in the movement of chemical species from areas of high concentration toward areas of lower concentration. Striven (1959) presented the mathematical treatment for the growth of a spherical vapor bubble in a superheated liquid when growth is controlled solely by the transport of heat and matter. He derived the parabolic growth law:

$$R = 2\beta(Dt)^{\frac{1}{2}} \tag{1}$$

where R is bubble radius, β is the growth rate constant, D is the effective bulk diffusion coefficient, and t is time. This investigation superseded an earlier study by Frank (1950) involving radian y s ymmetric phase growth controlled by diffusion.

Rosner and Epstein (1972) later derived the fundamental relationship describing the pressure of the gas in a growing bubble:

$$P_b = P_h + \frac{2\sigma}{R} + \rho_m\left(R\,R'' + \frac{3}{2}R'^2\right) + 4\mu\frac{R'}{R} \tag{2}$$

where P_b is the pressure in the bubble, P_h is the hydrostatic pressure, σ is surface tension, R is bubble radius, ρ_m is magma density, R' is the growth rate, R'' is the acceleration rate, and μ is the dynamic viscosity.

Sparks (1978) developed a numerical model for the case of a bubble in an infinite liquid based on empirical data to calculate the growth rate constant as well as poorly-defined values for the diffusion coefficient. Sparks was able to conclude that bubbles in magmas with high diffusivities and high water contents grow faster than bubbles in magmas with lower values for these parameters, and the bubbles attain a larger size upon reaching surface conditions during progressive decompression. He also concluded that faster magma ascent rates result in smaller bubbles because there is less time for growth. The study by Proussevitch et al. (1993) provided a more sophisticated numerical model, which invoked the solution of the full diffusion equations for bubble growth in a shell of melt of fixed volume. Their investigation did not incorporate any empirical values for key parameters, and their numerical approach examined instantaneous decompression under a range of parameter conditions. More recently, the numerical approach chosen by Toramaru (1995) involved a formulation to account for the viscosity effect on nucleation and bubble size distribution. He determined that there are regimes in which either diffusion or viscosity controls the nucleation and growth process. Barclay et al. (1995) presented improved analytical models for bubble growth during decompression of high viscosity magmas. Our

approach builds upon the experimental, theoretical, and numerical results of these previous investigators.

Experimental Method

Two high-silica rhyolite glasses were used as starting materials for high pressure hydration synthesis. Initial H_2O (total) contents were 0.17 ($\pm$0.06) wt% (Blank, 1993) and 0.78 ($\pm$0.01) wt% (Newman et al., 1986). Water contents and speciation were verified by FTIR (Fourier Transform Infrared Spectroscopy) analysis. Neither glass contained measurable CO_2. Cylinders of the natural glass were cored using a diamond drill from hand specimens and welded onto AgPd capsules. Capsules were held at 850°C for 10 to 14 days under hydrostatic pressures ranging from 700 to 1500 bars. Runs were quenched rapidly to room temperature with pressure held constant (Ihinger, 1991) producing optically clear but cracked glasses, Water contents and species were re-verified. High pressure hydration resulted in two glasses of 2.34 ($\pm$0. 10) and 4.35 ($\pm$0. 14) wt% average total dissolved H_2O with water contents homogeneous over 3 x 3 mm areas. A third glass, a sample with a natural initial H_2O (total) content of 0.78 ($\pm$0.01) wt% (Newman et al., 1986), was degassed at 1 atm without modification. In this report, we focus on experimental results and numerical modeling of the 2.34 wt% sample. We have reported previously on the experimental results for the 4.35 wt% sample (Mattioli et al., 1993; Mattioli and Stolper, 1994).

The bubble nucleation and growth rate experiments were conducted as follows. Irregularly shaped, doubly polished plates (-3 x 5 x 0.5 mm) of high-water content rhyolitic glasses were placed in a heating stage attached to a transmitted-light petrographic microscope. For the 1 atm exsolution experiments, samples were heated rapidly from 25°C to the target run temperatures between 500° and 700°C. Steady state temperature was generally achieved within 10 minutes of initial heating and controlled to $\pm$5°C of the target value throughout the run. Samples were continuously imaged using a CCD camera and recorded on an 8 mm VCR. Archived tapes were examined in detail a *postori* and selected frames were digitized at either 512 x 5'12 or 640 x 480 pixels with 256 gray levels. Digital images were then analyzed to determine individual bubble radii as a function of time since initial nucleation.

Numerical Theory and Approach

A computer model was formulated to simulate bubble growth in a magmatic system. Our approach to computer simulation is to begin with a simplified model and then enhance the model as physical parameters and processes are better understood, thereby providing a more realistic representati on 01 a complex natural system. These numerical simulations should allow us to quantify the contribution of other potentially important factors such as those related to viscous forces. Forward models allow direct comparison to experiments, where material properties such as the diffusion coefficient and bulk viscosity have been measured by other methods.

Our model applies a diffusion equation algorithm expressed in finite-difference form to operate upon a network of elementary cells representing a system of concentric shells. The system is designed to depict a center-stationary spherical gas bubble growing in the surrounding melt due to the transfer of a dilute component (dissolved H_2O) from the liquid to the gas (vapor) phase. The system is assumed to be isothermal and further obtains thermodynamic equilibrium at the gas-liquid interface. For the model presented in this report, we also assume that the gas phase within the bubble is homogeneous and that the viscosity of the melt is constant both at the gas/liquid interface and in the. far-field. We will report elsewhere on our refined model, which includes a coupled diffusion and water-concentration dependent viscosity as a function of radial distance from the bubble/melt interface (Hooper and Mattioli, 1996).

Bubble growth is induced by diffusion of the volatile component from the surrounding melt. This can be reduced to a case in which there is radial flow in a sphere. An analogous example is the flow of heat in a sphere, where the initial and surface conditions are such that the isothermal surfaces are concentric spheres (Carslaw and Jaeger, 1959). Therefore, radial diffusion with a constant diffusion coefficient can be described by the equation:

$$\frac{\partial C}{\partial t} = D\left(\frac{\partial^2 C}{\partial r^2} + \frac{2}{r}\frac{\partial C}{\partial r}\right) \tag{3}$$

where C is the concentration of the diffusing substance, D is the diffusion coefficient, r is the radial coordinate, and t is time. We can further substitute

$$u = Cr \tag{4}$$

such that Equation 3 can be simplified to the more manageable form:

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial r^2} \tag{5}$$

which is the linear diffusion equation in one dimension. The solutions to a variety of cases incorporating radial diffusive flow in a sphere can be found in Crank (1975).

For computer applications and a numerical solution, the diffusion equation can be transformed to a finite-difference form employing an array of concentration values in the bubble growth algorithm. In terms of the variables

$$R = r/a,\ T = Dt/a^2 \tag{6}$$

the non-dimensional equation for radial diffusion in a sphere of radius *a* is:

$$\frac{\partial c}{\partial T} = \frac{1}{R^2}\frac{\partial}{\partial R}\left(R^2\frac{\partial c}{\partial R}\right) = \frac{\partial^2 c}{\partial R^2} + \frac{2}{R}\frac{\partial c}{\partial R} \tag{7}$$

and after omitting the error term (Carslaw and Jaeger, 1959; Crank, 1975) the finite-difference approximation is

$$\frac{1}{R^2}\frac{\partial}{\partial R}\left(R^2\frac{\partial c}{\partial R}\right) = \frac{1}{i(\Delta R)^2}\left\{(i+1)c_{i+1,j}\ 2ic_{i,j}\ ^{+}\ (i\ \ 1)\ c_{i-1,j}\right\}, \qquad i \neq 1 \tag{8}$$

where $c_{i,j}$ is the concentration at the point $(i\Delta R, j\Delta T)$, AR is the spatial increment, and AT is the time increment. Both AR and AT must be kept sufficientl y small to prevent amplification of errors, i.e. oscillations in the solution.

The total flux of gas into the bubble is determined by:

$$4\pi a^2\left(D\frac{\Delta c}{\Delta R}\right)\Delta T \tag{9}$$

where a is the radius of the bubble, D is again the diffusion coefficient, Δc is the change in volatile concentration at the bubble-melt interface, AR is the spatial increment (same as the width of each concentric shell), and AT is again the time increment or time step (Crank, 1975).

Therefore, the numerical simulation models a system consisting of a sphere (the bubble) within an arbitrarily larger sphere (the melt). Based on computational efficiency, each simulation run divides the system into 10^3 to 10^6 individual cells, the linear representation of concentric shells. The algorithm keeps track of the position of the bubble edge within the system, the number of moles of H_2O within the bubble, the values of the various pressure terms (surface tension, viscous, inertial, excess, and total), bubble growth rate and acceleration, bubble volume, bubble halo width, and the ratio of halo radius to bubble radius. We define the halo as a volatile-depleted fringe that surrounds the bubble, which separates the homogeneous H_2O concentration within the bubble from a different homogeneous H_2O concentration within the melt in the far-field, as demonstrated in the diffusion profile (Fig. 1). An advantage of computer simulation is that it provides the methodology to rapidly treat a variety of cases, conditions, and the measured variation in physical parameters.

Numerical and Experimental Results

Rhyolite glass samples were heated rapidly to 500°C and bubble radii were measured for approximately three hours after initial nucleation (Fig. 2). Measured growth rates are nearly linear in contrast to that anticipated from the simple parabolic growth law (Eqn. 1) (Pérez and Mattioli, 1996). This deviation may be caused by viscous forces, which could inhibit early bubble growth. Viscous pressure is typically small compared with the supersaturation pressure, although it can be equal or greater than ambient pressure during the early stages of bubble growth. A greater understanding of these linear growth rates in the early phases of bubble growth is a subject of ongoing numerical and experimental investigations.

Experimental evidence indicates that bubble growth rates are not spatially uniform. Man y of the smaller bubbles had a lower growth rate than the larger ones, but even those bubbles that existed prior to heating (bubbles #1, 2, 3, and 8 in Table 1) did not necessarily have a higher growth rate. Although FTIR measurements demonstrated that initial water contents were homogeneous within error over 3 x 3 mm areas, perhaps small cracks in the sample or micron-scale heterogeneities in H_2O concentration are responsible for some of the experimental variability.

However, based on these results we can infer that the bubble surface area/volume ratio must exceed a critical value before it can overcome confining pressure and experience rapid growth.

Numerical simulations were run for a range of parameters relevant to bubble growth in a rhyolitic, high-silica melt. Parameter values include. a 500°C temperature, 2.34 wt% H_2O, 77.5 wt% SiO_2, volubility constant of 0.13 (derived from Sparks, 1978), and a viscosity of 3e+ 10 Pas (derived from Shaw, 1963; and Richet et al., 1996). Results from the simulations displayed similar growth rates to those obtained from the experimental results (Fig. 3). An initial period of relatively rapid growth was observed, which is not as clearly distinguishable in the majority of the experimental observations. Only bubble #3 (which existed within the sample prior to initial heating at 1 atm) showed a period of early rapid growth, whose rate slowed as a function of time. The comparable growth rates and initial bubble radii permit a simple interpolation to calculate a best-fit diffusion coefficient for each of the experimental bubbles (Table 1). Furthermore, the average of the values of these diffusion coefficients is within a half-order of magnitude, well within the range of error reported by Zhang et al. (1991), who conducted water dehydration experiments on rhyolitic glasses without causing their nucleation. Their experiments had initial water contents between 0.2 and 1.7% and were conducted between 400-5 50°C. Although our derived diffusion coefficients for the 2.34 wt% H_2O sample are within the 1 sigma errors reported by Zhang et al. (1991), we note that higher bulk water concentrations, such as our experiments on a 4.35 wt% H_2O-bearing glass require a higher bulk diffusion coefficient than that derived by extrapolation of the Zhang et al. (1.99 1) parameterization (Mattioli and Stolper, 1994).

Conclusion

Water-rich bubble growth has been observed and measured in real-time at 500°C and 1 atm total pressure. Numerical simulation with a diffusion-equation model provided comparable growth paths and permitted an inverse-solution approach to calculate the diffusion coefficient for each bubble analyzed in the experimental portion of the study. Our results confirm previous studies of bulk H_2O diffusion without phase transitions for rhyolitc glasses of intermediate water contents (e.g., Zhang et al., 1991).

Our current numerical modeling begins to examine the complex interactions in a multiple-bubble system and to parametrize the contribution of viscous forces. As we achieve a greater understanding of the degassing process from these bubble growth studies, we hope to improve our knowledge of those explosive eruptions for which exsolution of volatiles is the principal driving force.

References

Barclay, J., Riley, D. S., and Sparks, R. S.J., 1995. Analytical models for bubble growth during decompression of high viscosity magmas. Bulletin of Volcanology, v. 57:422-431.

Blank, J. G., 1993. An experimental investigation of the behavior of carbon dioxide in rhyolitic melt. PhD thesis, California Institute of Technology, Pasadena, California, 227 p.

Carslaw, H. S., and Jaeger, J.C., 1959. Conduction of Heat in Solids (2nd ed.). oxford University Press, Oxford and New York, 510 p.

Crank, J., 1975. The Mathematics of Diffusion (2nd ed.). Oxford University Press, Oxford and New York, 414 p.

Frank, F. C., 1950. Radially symmetric phase growth controlled by diffusion. Proceedings Royal Society London, v. 201 (Series A): 586-599.

Hooper, D.M., and Mattioli, G. S., 1996. Numerical simulation of H_2O bubble growth in rhyolite glasses. Journal of Volcanology and Geothermal Research, in preparation,

Ihinger, P. D., 1991. An experimental study of the interaction of water with granitic melt. PhD thesis, California Institute of Technology, Pasadena, California, 188 p.

Mattioli, G. S., and Stolper, E.M., 1994. Experimental constraints on volatile exsolution from rhyolite glass: implications for eruptive mechanism and volcanic degassing. Sixth Puerto Rico EPSCoR Annual Conference, San Juan, Abstracts with Program, p. 13.

Mattioli, G. S., Bashir, N., and Stolper, E.M., 1993. In situ, real-time determination of H_2O bubble formation and growth in rhyolitic glasses at 500" to 700°C: preliminary results and interpretations. 1993 WOVO Workshop: Volcano Observatories, Surveillance of Volcanoes and Prediction of Eruptions, Guadaloupe, Abstracts with Program, 3 p.

Newman, S.. Stolper, E.M., Epstein, S., 1986. Measurement of water in rhyolitic glasses: calibration of an infrared spectroscopic technique. American Mineralogist, v. 71:1527-1541.

Pérez, S. D., and Mattioli, G. S., 1996. Experimental determination of H_2O bubble nucleation and growth in water-rich rhyolitic glass. 31st American Junior Technical Meeting-1 6th Interdisciplinary Science Meeting of

Puerto Rico, University of Puerto Rico-Arecibo, Abstracts with Program, p. 110.
Prousevitch, A.A., Sahagian, D.L., and Anderson, A.T., 1993. Dynamics of diffusive bubble growth in magmas: isothermal case. Journal of Geophysical Research, v. 98:22,283-22,307.
Richet, P., Lejeune, A-M., Holtz, F., and Roux, J,, 1996. Water and the viscosity of andesite melts. Chemical Geology, in press.
Rosner, D. R., and Epstein, M., 1972. Effects of interface kinetics, capillarity and solute diffusion on bubble growth rates in highly supersaturated liquids. Chemical Engineering Science, v. 27:69-88.
Striven, L.E., 1959. On the dynamics of phase growth. Chemical Engineering Science, v. 10:1-13.
Shaw, H.R., 1963. Obsidian-H_2O viscosities at 1000 and 2000 bars in the temperature range 700° to 900°C. Journal of Geophysical Research, v. 68:6337-6343.
Sparks, R.S.J., 1978. The dynamics of bubble formation and growth in magmas: a review and analysis. Journal of Volcanology and Geothermal Research, v. 3:1-37.
Toramaru, A., 1995. Numerical study of nucleation and growth of bubbles in viscous magmas. Journal of Geophysical Research, v. 1.00:1913-1931.
Zhang, Y., Stolper, E.M., and Wasserburg, G.J., 1991. Diffusion of water in rhyolitic glasses. Geochimica et Cosmochimica Acts, v. 55:441-456.

Table 1. Calculated growth rates and best-fit diffusion coefficients (D) in the 2,34 wt% H_2O sample (at 500°C) for individual bubbles as averaged over the entire observation period of each bubble.

Bubble #	**dr/dt** (m/s)	Best-fit D (**m²/s**)
1	0.70e-09	3.60e-15
2	4.65e-09	1.55e-14
3	6.97e-09	2.07e-14
4	0.36e-09	2.27e-15
5	1.82e-09	7.28e-15
6	1.71e-09	6.95e-15
7	1 .34e-09	5.83e-15
8	3.75e-09	1.33e-14
9	2.49e-09	1 .02e-14
10	1.1 0e-09	5.07e-15
11	0.95e-09	4.52e-15
12	2.01e-09	7.86e-15
Average	2.32e-09	8.59e-15
Std. Dev.	1.92e-09	5.44e-15
Average (Zhang et al., 1991)		5.05e-14
Std. Dev. (Zhang et al., 1991)		1.83e-13

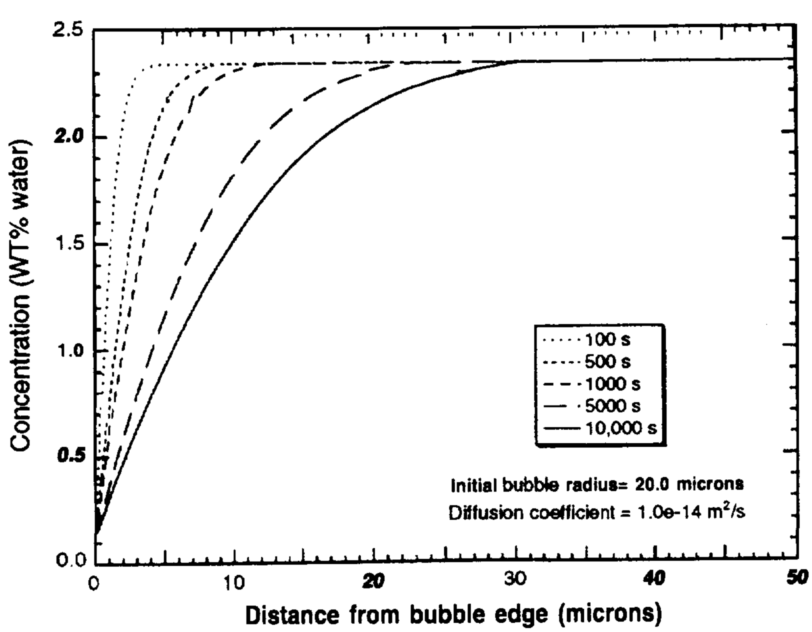


Fig. 1. Concentration profiles showing the growth of the diffusion halo with time.

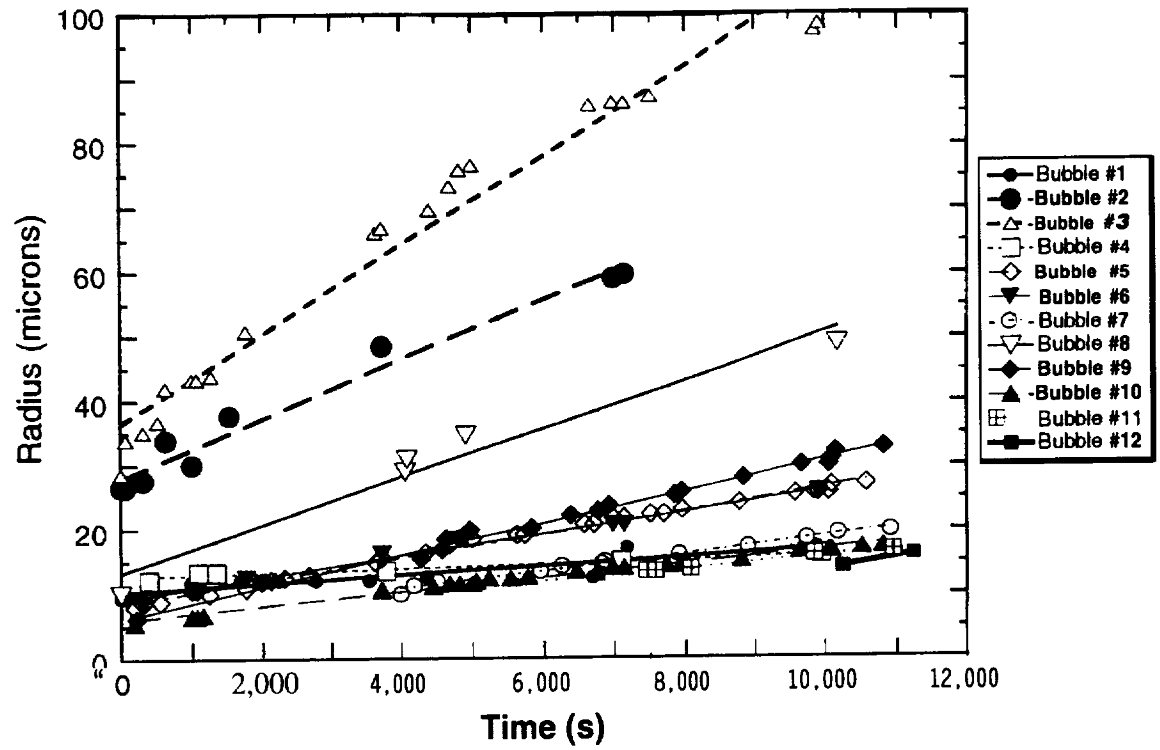


Fig. 2. Experimental] y-determined bubble growth rates from a 2.34 wt% H_2O rhyolite glass at 500°C. Each bubble is shown with its least-squares fit.

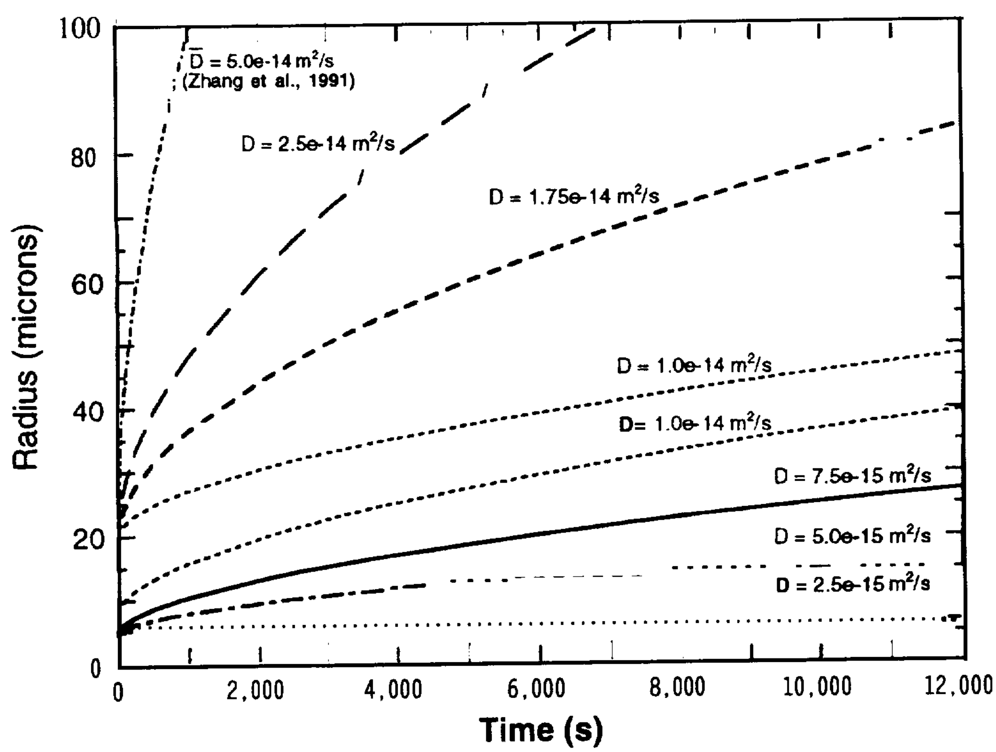


Fig. 3. Computer-simulated bubble growth rates run with different diffusion coefficients (D) and initial bubble radii to correlate with experimental results.

URC97059

Kinematic Flow Modeling and Computer Simulation for Volcanic Hazard Assessment at Soufriere Hills Volcano, Montserrat (B.W.I.)

Donald M. Hooper[1], Glen S. Mattioli[1], and Thomas P. Kover[2]

[1] Department of Geology
University of Puerto Rico
Mayagüez, Puerto Rico 00680

[2]GDE Systems, Inc.
P.O.Box **509008**
San Diego, California 92150

Introduction

The island of Montserrat lies in the British West Indies of the northern Lesser Antilles island arc. It is approximately 16 km long (N-S) by 9 km wide (E-W). Soufriere Hills volcano, located in the southern part of the island, consists of a series of andesitic lava domes. These domes are clustered near English's Crater, a horseshoe-shaped depression that is approximately 1 km in diameter and is open to the east-northeast. The origin and orientation of this feature has been attributed to a large sector collapse (Jansma and Mattioli, 1996) at approximate] y .4000 years B .P. (Smith et al., 1997). Chance's Peak dome, located on the western crater rim, is the highest (pre-eruption) point on the island at 915 m. Wadge and Isaacs (1988) summarized the available radiocarbon age determinations for the southern portion of the island, which indicate that the last period of major activity was between 16,000 and 24,000 years B.P. The last period of activity at Castle Peak dome, however, is more recent. The youngest pre-historic deposits, consisting of thin block-and-ash flow units, are restricted to the valleys draining the sector collapse. Radiocarbon ages for these deposits are between 750 and 200 years B .P. (Smith et al., 1997). Block-and-ash flow, surge, and mudflow deposits are dominant on the lower slopes of the Soufriere Hills Volcano (Rea, 1974; Wadge and Isaacs, 1988).

After several episodes of increased seismic activity over the past 100 years (Shepherd et al., 197 1), Soufriere Hills Volcano erupted for the first time in recorded history on 18 July 1995. Beginning with small phreatic eruptions, the overall strength and magnitude of the eruption has been escalating. Andesitic lava breached the surface in October 1995 and significant and dome growth began in November 1995. Dome collapse led to the generation of pyroclastic flows starting in April 1996 and continuing through the end of September 1996. These are relative] y gas-poor pyroclastic flows known as block-and-ash flows (Cas and Wright, 1987) and associated surges rather than gas-rich pumice and ash flows or ignimbrites associated with collapsing eruption columns (Smith and Mattioli, unpub. data).

Two general mechanisms can be proposed for the generation of these pyroclastic flows from lava domes: gravitational collapse or explosive collapse. Following the classification of Macdonald (1972), the gravitational mechanism has been described as a 'Merapi-type' pyroclastic flow caused by blocks falling from a collapsing dome, while the explosive mechanism has been described as a 'Peléan-type' pyroclastic flow caused by an explosion from the side of a dome. Observations by the field teams, an examination of the flow products, and computer simulations support a gravitational-collapse origin for most of the pyroclastic flows (Hooper et al., 1997). Because of their predominance at Soufriere Hills, an understanding of the pathways and kinematics of the flows generated by the disintegration of lava domes is critical for local hazard assessment.

Prior to the onset of the current eruption on Montserrat, Wadge and Isaacs (1988) used a digital elevation model (DEM) and a simple mathematical model of gravitational flow to create a map of the volcanic hazards from Soufriere Hills volcano. Their computer simulations followed the energy-cone model of Malin and Sheridan (1982). They produced a sequential hazard zone map that provided valuable information for emergency planning. Wadge (1996) later assessed the accuracy of their work and noted its usefulness as a guide for response early in the crisis, but the course of events was not accurately predicted and they made no attempt to assess the addition of a new lava dome. The Wadge and Isaacs model was designed more for a Plinian-type eruption or a "worst-case scenario" (Wadge, 1996) rather than the present activity, which has predominantly been pyroclastic flows generated by the gravitational collapse of the lava dome. The volcanic flow model we present here can be used to simulate several types of volcanic flows and has several significant differences from the Wadge and Isaacs model. Our model is highly adaptive for a variety of volcanic flows and incorporates parameters for basal friction, internal viscosity (or viscous drag), and inertial acceleration (or dissipation). We present some results from computer

simulations of pyroclastic flows and discuss the role of such computer models in hazard assessment.

Theoretical Background

McEwen and Malin (1989) created a two-dimensional flow model to be applied to a three-dimensional quadrilateral grid (U. S.G.S. digital elevation model or DEM). They described the total resistance to flow based on a model for the Mount St. Helens rockslide-avalanche suggested by Voight et al. (1983). The description was first derived by Mellor (1978) to describe snow avalanches and is in the form of the polynomial equation:

$$\tau = a_0 + a_1 \upsilon + a_2 \upsilon^2 \tag{1}$$

where τ is the resistance to flow, υ is velocity, and a_0, al, and a_2, respectively, can be related to internal and/or sliding friction, viscosity, and turbulence (inertia).

DTM Generation: The computer program used in this study has been called Flow3D or Model3D (Kover, 1995) and is based on a sliding block physical model with the inclusion of viscous and turbulent dissipation factors. The addition of the viscous and turbulent resistance parameters expands the capability of the model to include lahars and pyroclastic surges. This new program is similar to the McEwen and Malin (1989) model in its calculation of frictional resistance and its execution of flow calculations in two dimensions, but it differs in the type of DTM (digital terrain model) on which the flow model is applied and the way in which the viscous resistance is calculated. The McEwen and Malin model is applied to a grid-based DEM, while this program creates a DTM based on a triangular irregular network (TIN) which forms three-sided polygons from irregularly positioned three-dimensional triplet data. A TIN provides a favorable representation of a terrain surface for flow modeling because it allows the location of data points to be unrestrained and the vertices of the triangular mesh to precisely match the digitized data point values. The TIN can be constructed to be more or less detailed depending on topographic complexity. The Flow3D program then displays the terrain as a pseudo-three-dimensional surface that is shaded to highlight relief and colored to represent elevation.

The TIN is created by connecting the vertex points in such a way that the sides of the created triangle are as similar in length as possible. This is to prevent the creation of very long, narrow triangles, which detract from the accuracy and visual appearance of the DTM as well as the effective performance of any applied models. TIN creation used a Delauney triangulation method, which is loosely based on the technique described in Lee and Schachter (1980) and Jones et al. (1990). Delauney triangulation does not permit the circumcircle defined by the vertices of a given triangle to encompass any other vertices. If a triangle pair fails the circumcircle test, then swapping the shared edge of the triangle pair is required. The vertex data are stored as an array in which each data position holds the x, y, and z coordinates of a vertex.

Flow Models: The resistance due to friction is calculated using a variation of the Coulomb model similar to that employed by the energy line (Heim, 1932; Hsü, 1975; Sheridan, 1979) and sliding block (Pariseau and Voight, 1979) models:

$$\tau = g\mu \cos\theta \tag{2}$$

where τ is the frictional resistance, g is gravitational acceleration, μ is the coefficient of sliding or internal friction, and θ is the effective slope angle.

Flow3D uses a viscous dissipation algorithm to avoid problems associated with the calculation of the viscous resistance. This is the same approach employed in the model of Sheridan and Macías (1992). The viscous and turbulent resistances are calculated by multiplying the operator-defined viscous dissipation with the appropriate velocity at a fixed point in space and time. The path of a simulated flow is determined by vector arithmetic. The flow is modeled as an independent package sliding over an effectively smooth surface of varying slope. Mass, density, and volume are assumed to remain constant for each flow package. Therefore, erosion/deposition and the loss/gain of volatiles are ignored.

A variety of flow origin configuration schemes are available and include a single (point) flow source, multiple sources along an arc of a given radius, multiple random flow sources in a specified area, multiple flow sources in a line to simulate a fissure eruption, and multiple sources originating from the centers of all of the triangles within a specified area. Although each flow package behaves independently of the other flow packages, multiple flows initiated simultaneously are an efficient way in which to view the possible areal effects for a variety of flow scenarios. The

Flow3D program is fully described in Kover (1995),

Computer Simulation of Pyroclastic Flows

The Montserrat DTM was digitized from a 1 :25,000-scale topographic map published by the Government of the United Kingdom (Directorate of Overseas Surveys, 359/Edition 7, 1983). The chosen contour interval for digitizing was 250 ft. In order to assess the generation of pyroclastic flows from the new dome, we created synthetic topography based on the morphology from aerial and land-based photographs and field observations. For the simulations presented in this study, the maximum dome elevation was assumed to be 945 m (3100 ft.), which was the approximate height of the dome in August 1996 (MVO Daily and Scientific Reports, 1996). Our current numerical model is not yet capable of ingesting dynamic topography, although future modifications will incorporate this important modification.

In order to test the flow model, simulations were conducted on historic flows at Mount St. Helens (Washington, U. S. A.) and Unzen volcano (Kyushu, Japan). The simulations were compared in terms of flow path, flow termination, and, in the case of Mount St. Helens, flow velocity history. These simulations not only served as a test for the accuracy of the model, but also provided a relevant range of values for the empirical friction, viscous, and turbulent parameters.

On 7 August 1980, a pyroclastic flow originating from the crater dome at Mount St. Helens traveled north through the sector co]] apse left after the 18 May 1980 eruption. The flow was photographed by Hoblitt (1986) with a clock-equipped camera. The resulting series of photographs provides an accurate record of the path, runout, and velocity history for the flow front. Kover (1995) adjusted the friction, viscous, and turbulent parameters to match the characteristics of the actual flow and to obtain a relatively close fit. The simulated flow parameters are listed as MSH 1 in Table 1. The shorter runout distance of the simulated flow is most likely a manifestation of the actual pyroclastic flow inflating to a surge (Hoblitt, 1986; Kover, 1995).

The extrusion of viscous lavas near the summit edge of the Mt. Fugen (Fugendake) dome of Unzen volcano commenced on 20 May 1991 (e.g., Nakada and Fujii, 1993). Nakada and Fujii (1993) report that the continuous growth of lava domes and falls of lava blocks from their margins frequently generated pyroclastic flows (i e., Merapi-type). By matching the flow path and run out distance, Kover (1995) simulated the 8 June 1991 block-and-ash flow at Unzen volcano. His simulation (MU1 in Table 1) again obtained a reasonabl y close match with the actual flow.

These two sets of empirically-derived parameters were then used as input for the volcanic flow model with MONT1 corresponding to MSH1 and MONT2 corresponding to MU1 (Table 1). Numerous simulation runs were initiated with a variety of flow origin configuration schemes to model different initial flow directions. Although one flow per simulation run is possible, multiple flows initiated simultaneously along an arc were determined to be an efficient method of viewing subtle areal effects for a variety of flow scenarios. Because we are examining gravity-driven volcanic flows, each flow path had an initial velocity of O m/s.

Results from the simulations revealed that flow pathways quickly became channelized in the major drainages. After traveling down the Tar River valley, the flows entered the sea at the mouth of the Hot River. This is in agreement with the observation that the initial, gravity-driven pyroclastic flows had a mu-row areal extent and were confined to the major drainages leading from English's Crater. In the simulations using MONT 1, the flow velocities upon reaching the coast were 20-30 m/s, which is in accord with observed velocities of actual flows (M. Stasiuk, pers. comm.). Nakada and Fujii (1993) calculated the average velocities of four pyroclastic flows from the dome at Unzen Volcano and reported values between 14 and 28 m/s. In this same set of simulations, the flows reached a maximum velocity of approximately 50 m/s and traveled a maximum distance of 3050-3200 m before reaching the shoreline. The maximum duration of flow was between 110 and 130 s. The slight variation in values is a function of the number of flow paths, the path taken by each flow, and the flow origin configuration scheme. Figure 1 is a perspective view that shows a computer simulation of pyroclastic flows emanating from the new lava dome and descending the Tar River valley eastward to the shoreline. This simulation **uses** the MONT 1 parameters and multiple flow packages (16 flows) originating simultaneously from an arcuate configuration high on the dome. Although this figure is in black-and-white, the flow model is capable of color-coding both the elevation and flow velocity scales.

The simulations using the parameters from Unzen volcano (MONT2) generated flow paths that failed to reach the shoreline of the island. This result is consistent with the termination of

several of the smaller pyroclastic flows. This second parameter set created flows that reached a maximum velocity of about 33 m/s, traveled a maximum distance of roughly 2000 m, and had a maximum flow duration of approximately 120-125 s. A map view of one simulation is shown in Figure 2. The DTM has been geometrically corrected with north to the top of the figure. These gravity-driven flows were directed to the west and several of the 50 simultaneously-generated flows overtop the crater escarpment to the west and north. The capital city of Plymouth lies only 5 km to the west of the crater rim. This figure is significant because it illustrates that although the MONT2 parameters have shorter runout distances (due to greater resistance to flow) than the MONT1 parameters, they are still capable of overtopping the crater wall if the dome should collapse to the west. As the dome and talus apron continue to grow, it will become easier for volcanic flows to surmount topographic barriers.

The identification of a minor component of juvenile pumice detected in the late July 1996 pyroclastic flows (A. Smith, pers. comm.; MVO Scientific Reports) indicates the presence of an explosive component. By simply adding an initial velocity to the simulations, a first-order approximation of these conditions can be modeled. Adding an initial velocity to the flow pathways enables many of them to surmount the crater and valley escarpment and thereby endanger a greater portion of the island.

In addition to pyroclastic flows, the Flow3D program is also capable of modeling lahars (volcanic mudflows). Because they are intimately mixed with water, lahars quickly become confined to the main drainages in the simulations and have runout distances that generally reach the coast. Lahars can be generated from anywhere on the upper flanks of the volcano and their likelihood increases after a heavy or prolonged rainfall has had the opportunity to remobilize volcanic debris, especially ash.

The volcanic flow program also has an option to employ an energy-cone model (e.g., Malin and Sheridan, 1982). A density flow initiated at some elevation will move as potential energy is converted to kinetic energy minus friction. An energy line is the slope along which the frictional loss is balanced by conversion of potential to kinetic energy. The flow comes to rest where the energy line intersects the topographic surface. Although this is a more simplistic simulation, the application of an energy-cone model can be valuable for determining the flow paths of pyroclastic flows and surges. A more detailed discussion of lahars and the results from the energy-cone model will be presented in future publications.

Application to Hazard Assessment and Conclusion

Since the eruption began on 18 July 1995, there has been a gradual increase in eruption magnitude and destructiveness. English's Crater is open to the east-northeast and under the present eruptive conditions, the crater and the Tar River valley remain the areas most susceptible to various volcanic flow phenomena. This region and most of the southern portion of the island have wisel y been evacuated for a considerable period of time. To date, all pyroclastic flows have remained within 200 m of the confines of the crater walls and the Tar River escarpment, but significant additional dome growth or an increase in the magma ascent rate could lead to flows that overtop these boundaries. The potential for a volcano to enter a more destructive phase of eruption should never be underestimated. Less destructive surges have entered various canyons in different sectors of the volcano. Ash fall is dependent upon wind direction, but accumulating ash promotes the development of lahars, which may follow existing drainages leading from all sectors of the volcanic edifice. With relatively little effort by the users, the digital topography can even be updated to meet changing conditions.

A graphical simulation program like Flow3D is a useful tool for modeling future and historic volcanic flows. It is also a useful tool for illustrating complex volcanic processes. The three-dimensional representation of the energy relations of volcanic flows can be combined with digital terrain models of volcanoes to produce theoretical hazard maps. This methodology allows the scientist to begin to predict those areas most likely to be affected by a volcanic eruption, thereby helping to achieve the important goal of developing accurate models of volcanic processes so that the general public and authorities can be informed of the hazards. These hazards must then be clearly and efficiently communciated in a concise and timely manner. A graphical computer-simulation model represents a major step towards visualizing volcanic hazards and phenomena for both the scientist and the general public.

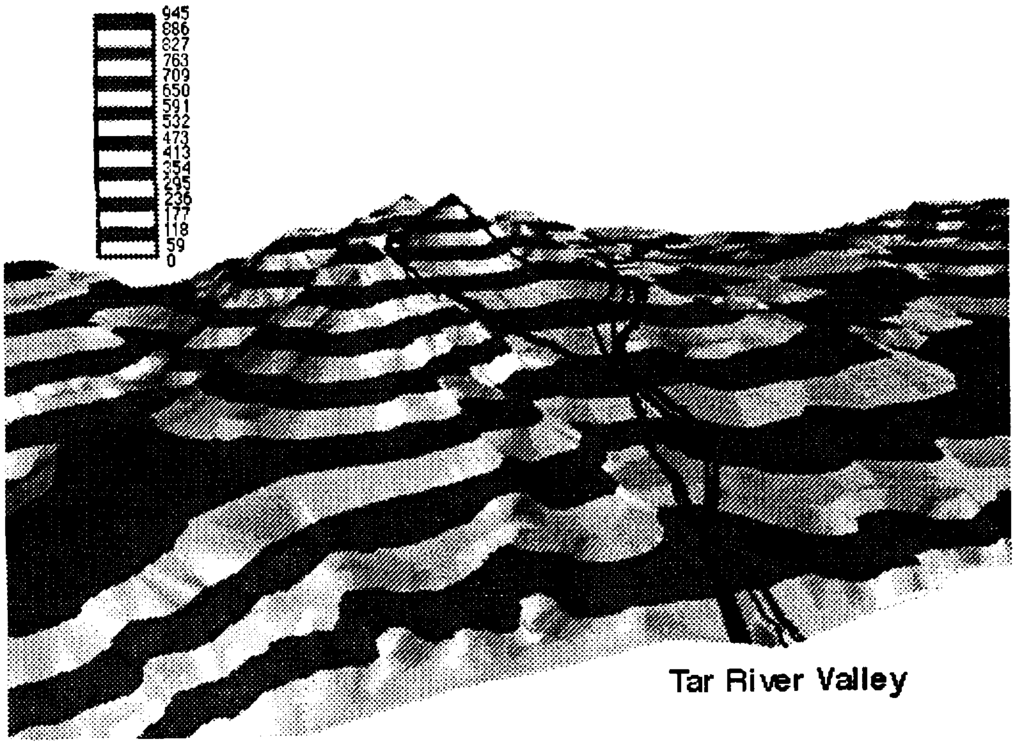


Fig. 1. Perspective view of simulated pyroclastic flows using MONT 1 parameters.

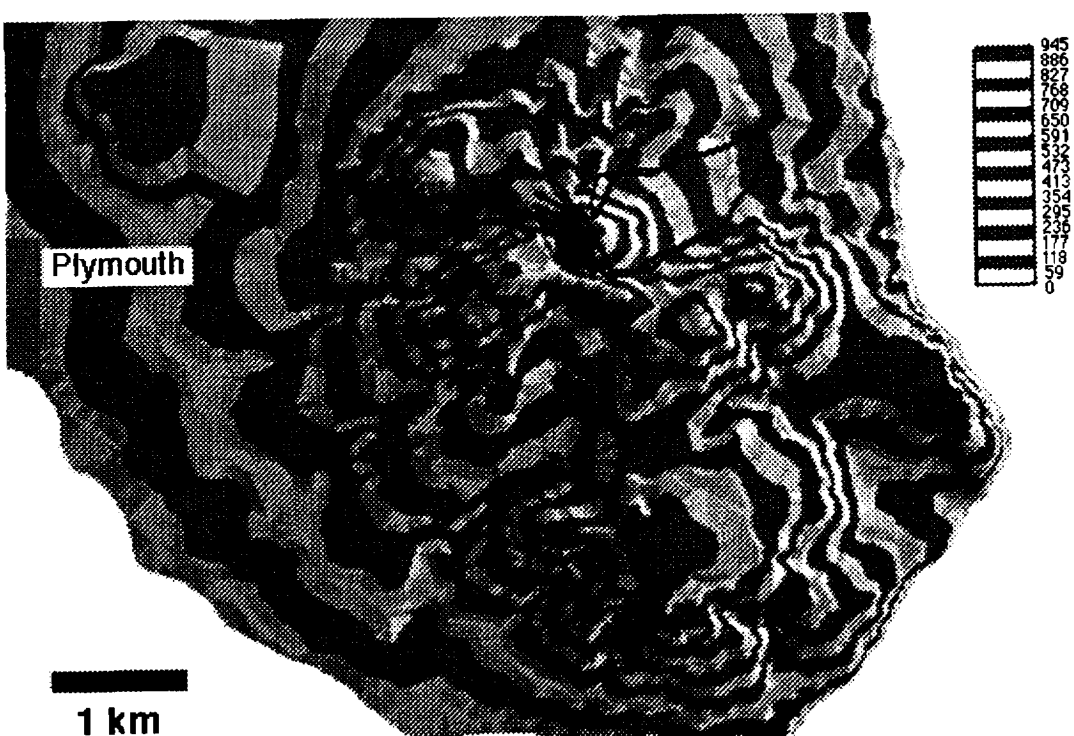


Fig 2. Map view of simulated pyroclastic flows using MONT2 parameters

Table 1. Flow model parameters.

Source or Simulation	Runout (km)	Frictional Coefficient (a_0)	Viscous Coefficient (al)	Turbulent Coefficient (a_2)
Mount St. Helens				
Hoblitt (1986)	5.7	0.19-0.23	–	–
MSH1(Kover, 1995)	4.7	0.10	0.01	0.001
MONT] (this study)		0.10	0.01	0.001
Mount Unzen				
Nakada and Fujii (1993)	5.5		–	
MU1(Kover, 1995)	~5.5	0.06	0	0.0125
MONT2 (this study)	--	0.06	0	0.01

References

Cas, R. A. F. and J. V. Wright. 1987. Volcanic successions, modern and ancient: A geological approach to processes, products and successions. Allen & Unwin, London.

Heim, A. 1932. Der Bergsturz von Elm. Deutsche Geologishe Gesellshaft Zeitschrift, v. 4, p. 74-115.

Hoblitt, R.P. 1986. Observations of the eruptions of Jul y 22 and August 7, 1980, at Mount St. Helens, Washington. U.S. Geological Survey Professional Paper 1335,44 pp.

Hooper, D. M., G.S. Mattioli, and T.P. Kover. 1997. Computer simulations of pyroclastic flows produced by gravitational dome collapse at Soufriere Hills, Montserrat, B.W.I. (Abstract) IAVCEI General Assembly, Puerto Vallarta, Mexico.

Hsü, K. 1975. Catastrophic debris streams (Sturzstroms) generated by rockfalls. Geol. Sot. Am. Bull., v. 86, p. 129-140.

Jansma, P.E. and G.S. Mattioli. 1996. Late Neogene dextral transtension in the northern Lesser Antilles island arc. In review, Tectonics.

Jones, N. L., S.G. Wright, and D.R. Maidment. 1990. Watershed delineation with triangle-based terrain models. J. of Hydraulic Eng., v. 116, p.1232- 1251.

Kover, T.P. 1995. Application of a digital terrain model for the modeling of volcanic flows: A tool for volcanic hazard determination. Unpublished M. SC. thesis, State University of New York at Buffalo, U.S.A.

Lee, D.T. and B.J. Schachter. 1980. Two algorithms for constructing a Delaunay Triangulation. Int. J. of Computer and Information Sciences, v. 9, p. 219-242.

Macdonald, G.A. 1972. Volcanoes. Prentice-Hall, Englewood Cliffs, New Jersey, U.S.A.

Malin, M.C., and M.F. Sheridan. 1982. Computer-assisted mapping of pyroclastic surges. Science, v. 217, p. 637-640.

McEwen, M.C. and M.C. Malin. 1989. Dynamics of Mount St. Helens' 1980 pyroclastic flows, rockslide-avalache, lahars, and blast. J. of Volcanol. and Geotherm. Res., v. 37, p. 205-231.

Mellor, M. 1978. Dynamics of snow avalanches. In: Voight, B. (ed), Rockslides and avalanches, 1, Natural phenomena. Elsevier, Amsterdam, p. 753-792.

Nakada, S. and T. Fujii. 1993. Preliminary report on the activity at Unzen Volcano (Japan), November 1990-November 1991: Dacite lava domes and pyroclastic flows. J. of Volcanol. and Geotherm. Res., v. 54, p. 319-333.

Pariseau, W.G. and B. Voight. 1979. Rockslides and avalanches -- basic principles and perspectives in realm of civil and mining operations. In: Voight, B. (ed), Rockslides and avalanches, 2, Engineering sites. Elsevier, Amsterdam, p. 1-92.

Rea, W.J. 1974. The volcanic geology and petrology of' Montserrat, West Indies. J. Geol. Sot. London, v. 130, p. 341-366.

Shepherd, J.B., J.F. Tomblin, and D.A. Woo. 1971. Volcano-seismic crisis in Montserrat, West Indies, 1966-67. Bull. Volcanol., v. 35, p. 143-163.

Sheridan, M.F. 1979. Emplacement of pyroclastic flows: A review. In: Chapin, C.E. and W.E. Elston (eds), Ash flow tuffs. Geol. Sot. Am. Special Paper 180, p. 125-136.

Sheridan, M.F. and J.L. Macías. 1992. PC software for 2-dimensional gravity-driven flows: Applications to the Colima and El Chichòn volcanoes, Mexico. (Abstract) 2nd Int. Volcanol. Meeting, Colima, Mexico, p. 5.

Smith, A. L., Roobol, M. J., Quiñones, E., and G. Mattioli. 1997. Volcanic history of Soufriere Hills, Montserrat. (Abstract) IAVCEI General Assembly, Puerto Vallarta, Mexico.

Voight, B., H. Glicken, R.J. Janda, and P.M. Douglas. 1983. Nature and mechanics of the Mount St. Helens rockslide-avalanche of 18 May 1980. Geotechnique, v. 33, p. 243-273.

Wadge, G. 1996. How did we do? Looking back on the 1986 assessment of volcanic hazards at Soufriere Hills. (Abstract) Science, Hazards and Hazard Management, The Second Caribbean Conference on Natural Hazards and Disasters, Kingston, Jamaica, p. 29.

Wadge, G. and M.C. Isaacs. 1988. Mapping the volcanic hazards from Soufriere Hills volcano, Montserrat, West Indies using an image processor. J. Geol. Sot, London. v. 145, p. 541-551.

URC97060

/// SA
S60 53

Non-Linear Post Processing Image Enhancement'

Shawn Hunt[†], Alex Lopez[‡], and Angel Torres[+]

'University of Puerto Rico Department of Electrical and Computer Engineering
Mayaguez, Puerto Rico, 00680
shawn@exodo.upr.clu.edu

*Easman Kodak Company
1700 Dewey Avenue
Rochester, New York 14650
lopez@image.Kodak.COM

"Motorola Inc.
Land Mobile Products Sector
8000 West Sunrise Blvd.
Ft. Lauderdale FL, 33322
torresan@plhp002.comm.mot.com

Abstract

A non-linear filter for image post processing based on the feedforward Neural Network topology is presented, This study was undertaken to investigate the usefulness of "smart' filters in image post processing. The filter has shown to be useful in recovering high frequencies, such as those lost during the JPEG compression-decompression process. The filtered images have a higher signal to noise ratio, and a higher perceived image quality. Simulation studies comparing the proposed filter with the optimum mean square non-linear filter, showing examples of the high frequency recovery, and the statistical properties of the filter are given,

1. Introduction

An artificial neural network (NN) based image post processor for image enhancement is presented. The goal of the research was to find a practical filter which would approximate the performance of the optimum mean square non-linear inter in recovering high frequencies. This has particular relevance for compressed images, and JPEG images were chosen for some of the simulation studies,

Many lossy compression algorithms reduce or completely eliminate high frequencies during the compression-decompression process. [n JPEG lossy compression for example, the spectrum of the image is quantized, and the high frequencies are usually rounded to zero. The basic problem here is then fundamentally different from the usual problem of removing noise from an image, The idea was to investigate to see whether there was information in the image that could be used to re-introduce high frequencies, and if we could implement a filter to do this. Linear filters cannot change parts of the spectrum that have been rounded to zero, and the non-linear filters available, both order statistic and others based on neural nets have, for the most part, been designed for the noise problem [1,2,3,4].

We have investigated a filter that improves the image both in terms of signal to noise ratio, and in perceived image quality. Quantization to both zero and non-zero values will cause image degradation. While linear filters may improve perceived quality caused by quantization to non-zero values, they can do nothing for values quantized to zero, Quantization to non-zero values can make the image look grainy, and linear smoothing filters can be used to improve the perceived image quality. The signal to noise ratio however, will decrease. This is acceptable, since most JPEG images are ultimately used for viewing, and perceived quality is considered to be the standard, while signal to noise ratio is used only as a secondary measure, Linear filters can only increase or decrease spectral components, and so wi II have no effect on frequency components rounded to zero. This is where the non-linear filters come into play. We have called these 'smart' filters because they must 'decide' when high frequencies have been lost, and re-introduce them. These decisions are statistically based, and must be learned during training, The first part of the research compares the optimum non-linear filter in the mean square sence to the neural filter, the the neural filter is applied to a real world problem, JPEG images.

2. JPEG Image Compression

Without going into the details, here are the relevant aspects of the JPEG lossy standard. It might be usefu I to distinguish here between two variations of JPEG, the lossy and the lossless types. They are fundamentally different, and are implemented with completely different algorithms. We are interested in the lossy type here.

[1] Acknowledgments: This work was supported by NASA under Grant No. CAN NCCN-0088, the National Science Foundation under Grant No. CDA-89 13486, and the Industrial Affiliates Program at UPR.

The JPEG lossy standard proposes the following compression and decompression steps. An image is first segmented into blocks of $8x8$ pixels. 2^{b-1} is subtracted from each pixel, where b is the number of bits per pixel. The dc value is calculated and compared to the dc value of the previous block. The difference between them is stored. A discrete cosine transform (DCF) is then performed on each block. The results of the DCT are then quantized using a 64 element quantization table. The values from the look up table are then coded using either arithmetic or Huffman coding. For decompression, the quantization table values are recovered from their coded values, the look up table is used to regain the quantized values, and an inverse DCT is performed.

As a result of the quantization, small values of the DCT are rounded to zero. This can manifest itself visually differently depending on the amount of quantization. An image generally has more energy in the lower frequencies than in the higher ones, which means that the higher frequency components of the DCT are usually the ones quantized to zero. A loss of sharp edges, or blurring, can be seen in the compressed-decompressed image. We are concerned with non-linear titers which can be used to remove some of the blurring effects.

3. Neural Filter Topology and Training

A feedforward neural net topology shown in figure 1 was chosen for the filter. Both one and two hidden layers were used in the simulations. Since they gave very similar results, we have presented only the one hidden layer case here. The number of hidden layer nodes was varied from 3 to 10.

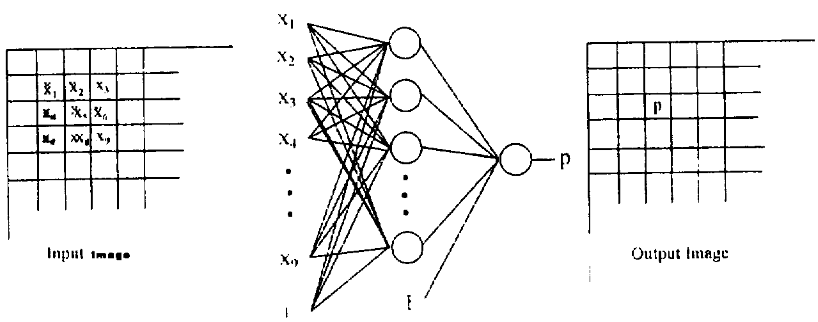


Figure i. Topology of filter showing input and output.

The images have intensity values ranging from 0 to 2^b-1, where b is the number of bits per pixel. The images used in the simulations were 8 bits per pixel, or intensity values ranging from 0 to 255. These values were scaled in order to have inputs and outputs between 0 and 1.

Training patterns were produced using the JPEG compressed-decompressed image as input, and the original image as the target values. Let the number of rows and columns in the image be R and C respectively. Let $p(i,j)$ be the scaled pixel in row i, column,) of the original image, where $1 \le i \le C$, and $1 \le j \le R$. Also, let $p_{JPEG}(i,j)$ denote the scaled pixel in row i, column j of the JPEG processed image. Each training pattern consists of nine input values,

$$p_{JPEG}(i,j) : k-1 \le i \le k+1,\ m-1 \le j \le m+1$$

and one output value,

$$p(k,m).$$

Since the inputs correspond to a pixel and its eight nearest neighbors, pixels on the borders were not used as training patterns, The method for selecting training patterns was different for the different experiments, and is described in the results section. Supervised training was done with both the Back-propagation and Levenberg-Marquardt algorithm. The Levenberg-Marquardt algorithm proved superior, and the results shown use this training method.

Once training was completed, filtering was done using a 3 by 3 pixel sliding window as shown in figure 2. All pixels except those on the border were filtered. 'Thus, the filtered images were 2 rows and 2 columns smaller than the original.

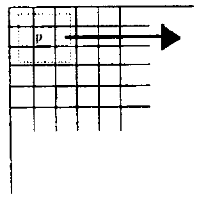


Figure 2. Filtering process using a sliding window.

4. Experimental Results

The non-linear filters proposed use the mean square error as a performance measure. The optimum filter would output the conditional expected value of the output pixel given the input pixels,

$$p=E\{ x_5 \mid x_i, i=1,2, . 9\}$$

where p and xi are as in figure 1. This can be estimated from the images to be filtered, but cannot be implemented for any but the simplest filters. For even the relatively small 3 by 3 pixel filter and 8 bit images used here, there are 256^9 conditional expected values. Computing and storing these is not possible on most computing platforms. Anything over two pixels becomes computationally too expensive. The first experiment therefore uses two pixels as input. Various images were processed, using both the optimum filter and the neural filter. The high frequencies of the images were removed to produce an average signal to noise ratio of 8.2 dB relative to the original image. The complete images were used to produce the conditional expected values for the optimum filter. After filtering the average signal to noise ratio incresed to 11.72 dB. The neural filter was trained using only 1/100 of the images, or 1/ 100 of the information used by the optimum filter. Still, the average signal to noise ratio increased to 10.1 dB after filtering. These results encouraged the use of higher order filters.

The next simulation done was to test whether a non-linear filter of this type could restore the high frequencies of a JPEG compressed image. An image was selected, and every 100^{th} pixel was used as a training pattern. Once the training was complete, the filter was used on the same JPEG image. Figure 3 shows the results,

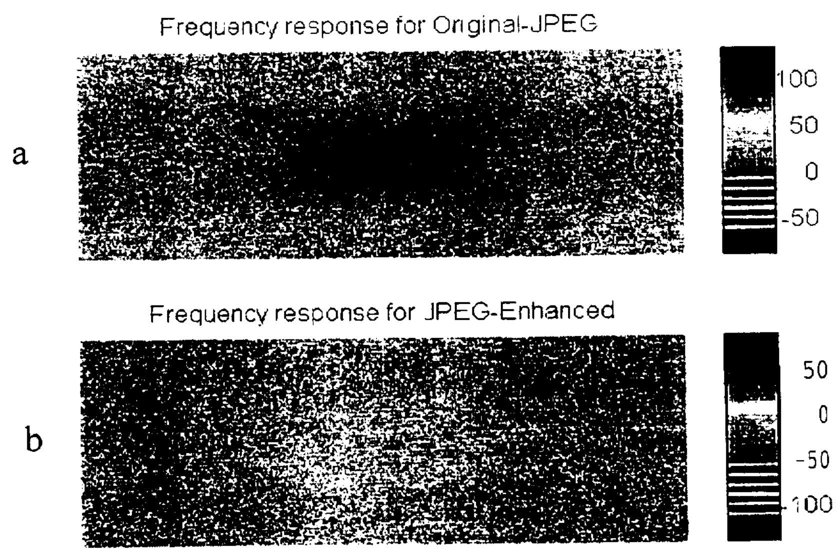


Figure 3. a) Difference of magnitude of FFT of original image and JPEG image, b) Difference of magnitude of FFT of filtered image and JPEG image

Figure 3 a is the difference between the magnitudes of the FFT of the original image, and the FFT of the JPEG compressed-decompressed image. The plot shows large differences at the high frequencies, indicating that these have been lost. The filtered image is compared to the JPEG image in figure 3 b. Again the differences are in the high frequencies, indicating that the filter did indeed introduce high frequencies. These high frequencies were not just introduced randomly, and the filtered image was better visually, and had a higher signal to noise ratio,

These results, although encouraging, are not general, since the training patterns came from the image to be filtered. A second filter was trained using training patterns from various different images. This filter was then used to filter three different images. Results of the output signal to noise ratios are shown in Table 1

Table 1. Signal to noise ratio in dB of the images before and after filtering

	JPEG Image	Filtered Image
Image 1	29.44	30.06
Image 2	20.86	21.42
Image 3	21.94	23.18

As can be seen, the signal to noise ratio increased for all three images after filtering. More importantly, the images improved visually, In order to appreciate the difference between the images, a magnified portion of one of the images is presented in figure 4. It is clear how the JPEG compressed image is blurred compared to the original. [n figure 4 c. it can be seen that the image after filtering has edges that are more clearly defined.

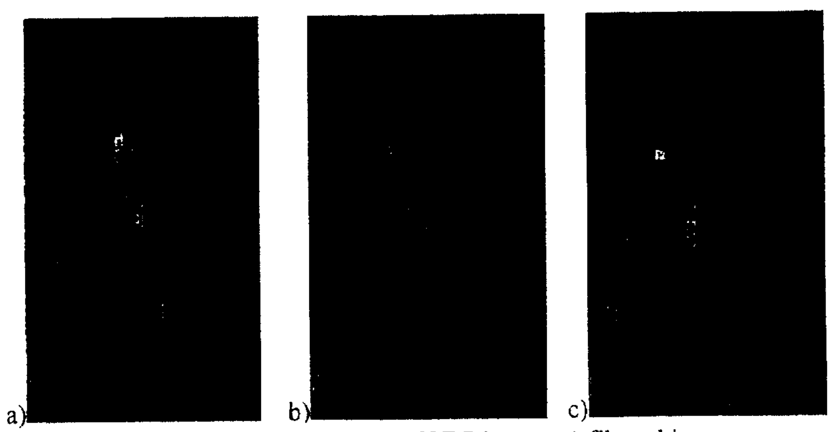


Figure 4. a) original image, b) JPEG image, c) filtered image,

The frequency plots presented above give us an idea of what the filter is doing, but the filter is not linear, so frequency response plots do not specify the filter completely. We used statistical plots in order to get a clearer picture of what the filter was accomplishing. We used two different plots for this. The first are the histograms of the original image, the image after JPEG compression, and after filtering. The plots for one of the images is shown in figure 5. The removal of the high frequencies can be seen to have the effect of smoothing the histogram.

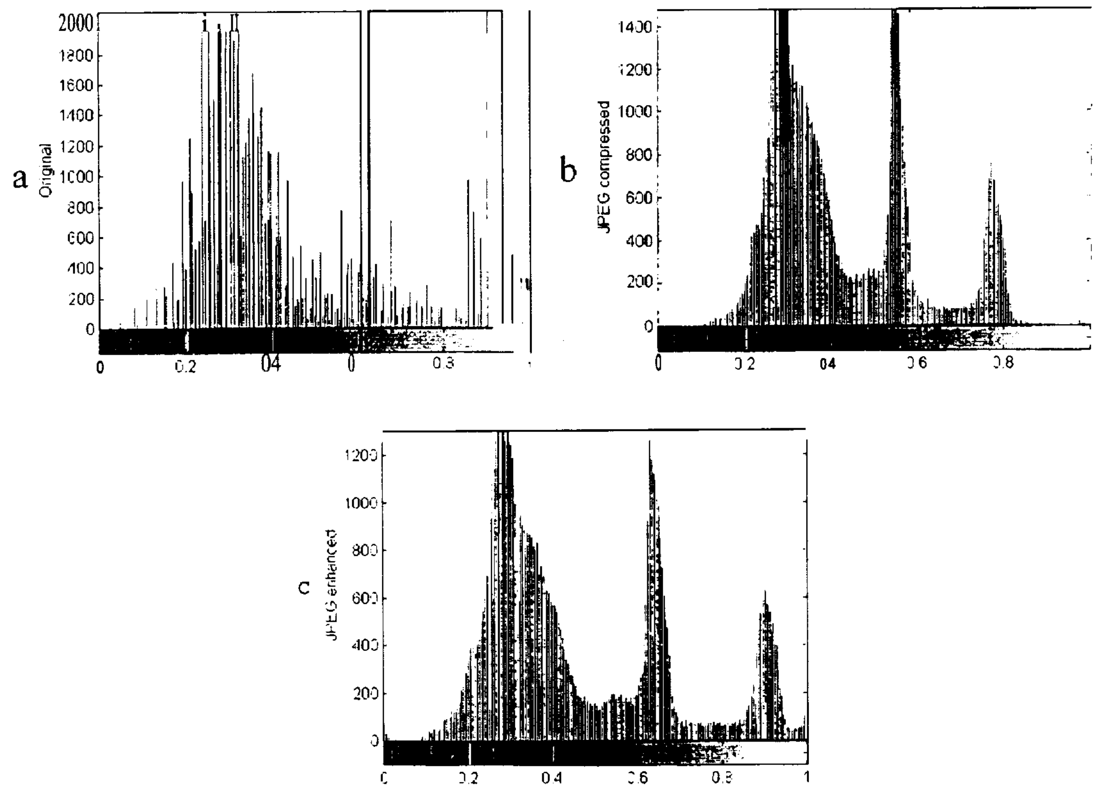


Figure 5. Histograms of a) original image, b) JPEG image c) filtered image.

The second plots were the pixel intensities before filtering vs. the pixel intensities after filtering. This will show exactly how the filter is modifying pixel intensities. If no changes had been done, then we wou ld see only a line through the origin with slope 1. A plot for one of the images is shown in figure 6. Here it can be seen that the filter does not radically change the image, and that more mollification is done on the low pixel intensities than on the high.

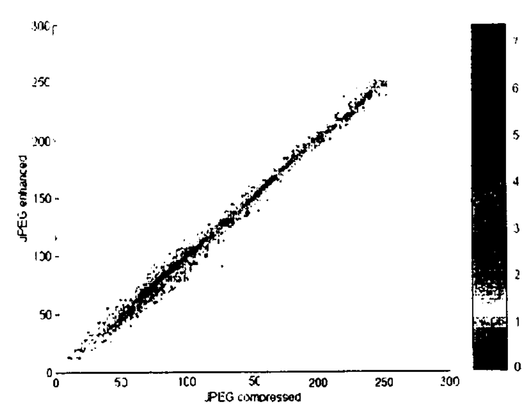


Figure 6. Plot of pixel values of the JPEG image vs. the JPEG enhanced filtered image.

5. Conclusions and Future Work

A non-linear filter, based on the feedforward NN topology was presented. This filter has shown to be useful in recovering high frequencies lost during the JPEG compression-decompression process, Designing filters to increase the signal to noise ratio of two dimensional signals by introducing lost high frequency components is useful in may different image processing applications. Further work needs to be done testing the filter on a wider image database, and in refining the training patterns to enhance performance.

References

[1] H. Hanek, N. Ansari, Z.Z. Zhang, " Comparative study of the generalized adaptive neural filter with other non linear filters." *IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 1*, p. 649-52, April 1993.

[2] 1. Pitas, "State of the art morphological and nonlinear digital image processing, " *IEEE Colloquium on 'Morphological and Nonlinear Image Processing Techniques'*, June 1993.

[3] A. Marston, S.-K Park, " Nonlinear filter design using artificial neural networks," *IEEE Proceedings of SOUTHEASTCON '91*, vol.2, pp. 931-4, April 1991.

[4] T. Fechner, " Nonlinear noise filtering with neural networks: comparison with Weiner optimal filtering," *Third International Conference on Artificial Neural Networks*, p.143-7, May 1993.

URC97061

BARIUM STRONTIUM TITANATE THIN FILMS FOR DRAM APPLICATIONS

P. Jana and R. K. Pandey
Center for Electronic Materials
Department of Electrical Engineering
Texas A&M University
College Station, TX 77843-3253.

I. INTRODUCTION

The progress made in silicon based semiconductor and ULSI technology is mainly driven by dynamic random access memories (DRAM) and in part by high speed RISC processors. It is interesting to note that DRAM has always been driving microlithography and front half technology such as isolation and Field Effect Transistors (FET) whereas RISC processors have driven the back half technology such as multi-level interconnection, metalization and planarization. The DRAM technology has always been advancing at a steady rate, producing one new generation every three years and it is likely to continue unabated for several generations more. [1] Each generation offers 4x increase in density, 1.4x increase in chip size with 3x reduction in cell area and 1.4x reduction in feature size. The market for DRAM is also progressing at a similar rate. This demand has been the driving force behind the advance of the semiconductor processing technology.

A typical DRAM cell consists of a capacitor for charge storage and an MOSFET for charge switching. [2] This configuration is called 1-Transistor 1-Capacitor (1 T- 1 C) and is very popular owing to its simplicity and small size. The cell is accessed by turning on the MOSFET via the word line. Charge is then transferred into/out of the capacitor, The information stored in the cell is read destructively and subsequently restored by the sense amplifier. Fig. 1 shows a 1 T- 1 C DRAM cell.

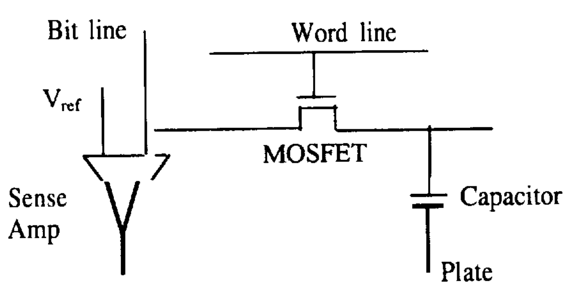


Fig. 1. Schematic of a 1T-1 C DRAM cell

With the increase in packing density and reduction in feature size the capacitor is taking up most of the semiconductor real estate on the integrated circuit. Capacitance offered by a cell of effective area A and effective dielectric thickness t is given by the formula

$$C = \varepsilon_0\varepsilon_r\,(A/t) \qquad (1)$$

where ε_r is the relative permittivity or the dielectric constant of the material. Currently used material is a combination of SiO_2 and Si_3N_4. [3] This layer is called as Oxy-Nitride (ON) or Oxy-Nitride-Oxide (ONO) and has a dielectric constant of about 6. It is estimated that the minimum capacitance that the cell should offer for the satisfactory storage of charge and operation of the DRAM cell is 30 fF/μm^2. [4] This value is called the Critical Capacitance C_{crit}. As the cell size reduces the dielectric is deposited or patterned into complex non-planar (3D) structures in order to maintain the required capacitance. However, for the new 1 Gb and higher density memory it is not possible to realize a compact DRAM cell using these conventional materials because of the extremely large capacitor area required. The need for denser memories has thus placed a greater constraint on the materials employed. Barium Strontium Titanate ($Ba_{1-x}Sr_xTiO_3$) has been projected to be an effective replacement and a material of choice for future generations of DRAM. The other leading possible materials include strontium titanate and lead zirconium titanate in both doped and undoped forms. However the inherent disadvantages associated with these materials have put them behind BST in the choice of high permittivity material. [5]

II. THIN FILM DEPOSITION

Thin film deposition of BST has been pursued by various techniques including sputtering [6] [7] [8], chemical vapor deposition [9] [10] [11] and sol-gel [12] methods. We have chosen to use the low cost process of Metal Organic Decomposition (MOD) which includes the preparation of organometallic precursors and subsequently deposit them by the spin-on technique. The selected composition is $Ba_{0.5}Sr_{0.5}TiO_3$. The MOD is a technique for producing inorganic thin films without processing in vacuum or going through a powder step. [13] A metal-organic compound has a metal atom bonded to an organic radical through a hetero-atom like oxygen, nitrogen, sulfur or phosphorous. The individual organometallic compounds are mixed in cation stoichiometry to form a true solution. This solution is then deposited on a substrate by any suitable method. The substrate is then heated to pyrolyze the formulation thereby getting rid of the solvent and then to decompose the organometallic compound to produce an inorganic film. A further heating step apropos the annealing is required to control the physical properties including stoichiometry. Multiple processing is done in order to achieve a greater thickness, prior to annealing. Some of the significant advantages of MOD processing over alternate techniques are the yield of equilibrium phases (of desired systems) at relatively low temperatures, fine grain polycrystalline films and relatively high stoichiometry. Also, the cost of production is considerably less than most methods of film growth. The disadvantage is the possible cracks and pinholes that can appear during the thermal processing, due to the rapid change in material volume.

The MOD chemicals are very sensitive to moisture, particularly for long shelf life. In order to circumvent the elaborate moisture removal procedures water soluble compounds are chosen to form the precursors. [14] Acetates are known to be reasonably good candidates for this purpose. We chose barium acetate and strontium acetate and for the titanium atom requirement, a chelated liquid solution of titanium acetyl acetonate commercially available as Tyzor GBA, a proprietary chemical manufactured and marketed by Du Pent.

Required molar quantities of barium and strontium acetates are weighed and saturated solutions are prepared in water. From the data sheets, volume of Tyzor GBA to be pipetted for the required molar quantity of titanium is determined. Methanol is the common solvent for the final precursor solution. However both barium and strontium acetates are insoluble in methanol.

A little quantity of water is added to titanium acetyl acetonate. In the presence of this additional water and some methanol all the cation sources mix well to form a homogeneous liquid. A slightly acidic pH is maintained (about 6.5). The quantity of methanol is primarily decided by the required viscosity of the solution. The solution is stored in air tight bottle and it has reasonably long shelf life of several months.

Platinized silicon wafer is used as the substrate. P type (100) silicon wafer with e-beam evaporated 1000 Angstroms of platinum and 400 Angstroms of barrier layer were purchased from commercial vendors. The wafer is cut into small pieces of about 1 cm x 1 cm dimension. They are cleaned to be free of all grease, oils and inorganic contaminants with the aid of ultrasonic bath and standard cleaning procedures. The first step in cleaning is to wash the substrate with water and detergent to remove inorganic and surface adherents. They are then dipped in acetone to remove any grease on the surface. Cleaning with isopropyl alcohol removes the acetone from the wafer. Methanol dissolves and removes any remaining organic impurities and solvents. The substrates are then dipped once again in the ultrasonic bath, this time with distilled water only. After blow drying the wafer front with pure nitrogen gas they are baked in a low temperature (-150 'C) oven for a short duration. A 4" wafer spin processor (model: 4 NPP; Lauren Technologies) is used for spin coating the deposition. The substrate is mounted in the spinner; a few drops of viscosity controlled stock solution dropped on it using a simple liquid dropper and the wafer spun at a suitable speed ranging between 1000 and 3000 rpm. The coated substrate is pyrolized for a few minutes at about 150 ^{o}C to remove organic solvents and water. The process is repeated to deposit multiple layers. Individual layers are too thin to be analyzed and has a lot of cracks and pinholes. Finally the film is annealed at a higher temperature between 500 ^{o}C and 800 'C for about 30 minutes. This promotes crystallization of the film. A flow diagram for the deposition process is shown in Fig. 2.

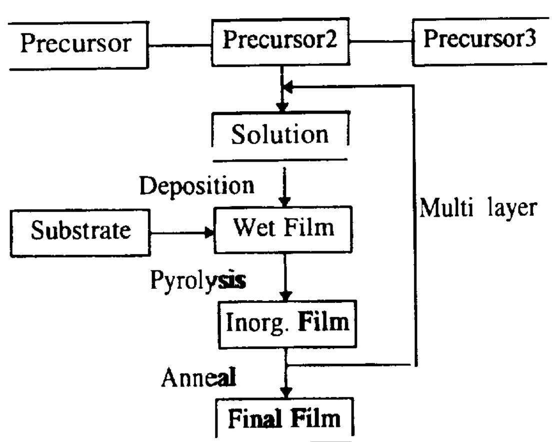


Fig. 2: Flow diagram for MOD process

The films produced are then coated with evaporated gold contacts for the purpose of electrical measurement.

III. RESULTS AND DISCUSSION

A schematic of the film structure is given in Fig. 3. The first films deposited were on platinized silicon with a TiO_2 barrier layer. This caused a problem in identifying the composition. Wavelength Dispersive Spectroscopy (WDS) produced results that were inconclusive regarding the contents of titanium in the film since the x-ray counts conflicted with those generated from the barrier layer. Thus wafers with tantalum or zirconium oxide barrier layers were used for the first films. The WDS is a bulk sample analysis program. In order to account for the film analysis "GMRfilm" program is used. This program compensates for the bulk effects and provides a quantitative analysis based on the underlying (substrate) layers and their composition too. Since this is a complex structure involving quaternary compound for the film and multiple layers for the substrate there was a need to determine the reliability of the measurement from the electron microprobe. Thus a few films were deposited on sapphire (Al_2O_3) substrates. Upon comparing the results obtained for the different films it was found that the system was indeed precise and the readings very repeatable. We obtained the desired composition of $Ba_{0.5}Sr_{0.5}TiO_3$ in the final film. The composition was accurate to 0.01 for the barium and strontium atoms.

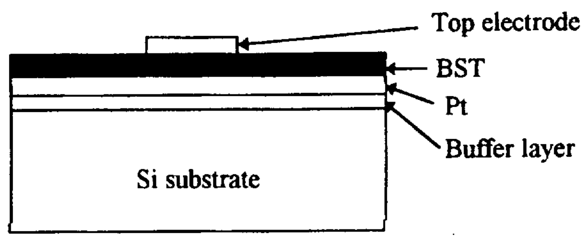


Fig. 3: Deposited film schematic

Individual film layers deposited by the spin-on process are extremely thin. Each layer was found to be of the order of 100 Angstroms thick. This causes a problem in obtaining the x-ray powder diffraction pattern. The presence of a large number of pin holes only compounds this problem. Upon deposition of about 10 layers it was possible to obtain a pattern showing the presence of weak lines of diffraction. However the weak lines and insufficient crystallization showed the need for further tuning of the annealing parameters.

Spinner speed was varied from about 1000 rpm to about 3500 rpm and spinning duration of 15 seconds was found adequate. Viscosity of the film was suitably adjusted. It was found that lower speeds produce thicker films as expected. But these films also contained more cracks since the process of pyrolysis left large changes in volume of the deposited matter. Thus optimized speed is in the neighborhood of 3000 rpm. Upon exposing the solution to air for a longer duration methanol evaporation caused the stock to become more viscous. Thus care has to be taken to maintain viscosity as well as operating conditions. Annealing at temperatures of about 700 ^{o}C was found to be in the right range. Annealing at under 550 ^{o}C failed to produce any detectable BST phase suggesting that the higher temperature played the part. Extended duration did not have any effect on the quality of film.

Detailed electrical measurements and studies are to be made. Preliminary studies showed that the films are continuous and electrically insulating. Dielectric constant, leakage current and other parameters will be determined. The film is not expected to be ferroelectric since the MOD

process produces fine grain structure of BST and the long range ordering required for ferroelectricity and domain dynamics cease to exist.

IV. CONCLUSION

BST thin films of composition $Ba_{0.5}Sr_{0.5}TiO_3$ have been deposited using MOD spin-on technique. It can be seen that the metal organic decomposition spin-on process is a very highly reliable and accurate film deposition technique with composition as dictated by the precursor solution stoichiometry. Annealing plays a very decisive role in the crystallization of the required phase and producing the final film. Multiple coatings are a necessity in order to obtain any workable thickness and integral film.

V. REFERENCES

1. B. E. Gnade, S. R. Summerfelt and D. Crenshaw; "Science and Technology of Electroceramic Thin Films" eds :0. Auciello and R. Waser; Kluwer Academic Publishers (1995) p. 373.
2. Digital Design: Principles and Practices; 2nd Edition; Prentice Hall 1994 p. 754.
3. D. E. Kotecki; ISIF '96; to be published in Integrated Ferroelectrics.
4. P. C. Fazan; Integrated Ferroelectrics 4 (1994) p. 247.
5. K Koyama, T. Sakuma, S. Yamamichi, H. Watanabe, H. Aoki, S. Ohya, Y. Miyasaka and T. Kikkawa; IEDM '91 (1991) p. 823.
6. S Y. Cha, S. H. Lee and H. C. Lee; presented at ISIF '96, To be published in Integrated Ferroelectrics.
7. S Ohfuji, M. Itsumi and H. Akiya; presented at ISIF '96, To be published in Integrated Ferroelectrics.
8. C. J. Peng, H. Hu and S. B. Krupanidhi; Proceedings of the 8th IEEE-ISAF '92, 1992 p305.
9. C. Basceri, S. K. Streiffer, A. I. Kingon, S. Bilodeau, R. Carl, P. C. van Buskirk, S. R. Summerfelt, P. McIntyre and R. Waser; presented at MRS Spring '96, To be published.
10. P. Kirlin, S. Bilodeau and P. C. van Buskirk; Integrated Ferroelectrics 7(1-4) 1995 p307.
11. M. Yamamuka, T. Kawahara, T. Makita, A. Yuuki and K. One; Japanese Journal of Applied Physics, 351996 p729.
12. D. Tahan, a. Safari and L. C. Klein; Proceedings of the 9th ISAF '94 (1994) p.
13. R. W. Vest, G. M. Vest, A. S. Shaikh and G. L. Liedl; "Metallo-Organic Decomposition Process For Dielectric Films" Annual Report for Office of Naval Research; Purdue University 1988.
14. G. H. Haertling; Ferroelectrics 116 (1991) p. 51.

Page intentionally left blank

URC97062

A Numerical Simulation of a Normal Sonic Jet into a Hypersonic Cross-flow

Damon K. Jeffries, Ramesh Krishnamurthy, and Suresh Chandra
NASA Center for Aerospace Research, College of Engineering
North Carolina A & T State University, Greensboro NC 27411

36 /34

Abstract

This study involves numerical modeling of a normal sonic jet injection into a hypersonic cross-flow. The numerical code used for simulation is GASP (General Aerodynamic Simulation Program.) First the numerical predictions are compared with well established solutions for compressible laminar flow. Then comparisons are made with non-injection test case measurements of surface pressure distributions. Good agreement with the measurements is observed. Currently comparisons are underway with the injection case. All the experimental data were generated at the Southampton University Light Piston Isentropic Compression Tube.

Introduction

A jet in cross-flow (JICF) consists of a jet exhausting at a large angle into a freestream flow. It is a flow field which is relevant to a wide variety of technologies and applications. When the primary importance is the mixing of the jet with the cross-flow, then an extensive application could involve heat transfer, gas turbines, and fuel injection. The JICF also plays a crucial role in technologies such as vehicle attitude control, the hover in-ground effect (HIGE) of some military aircraft, and the vertical and/or short takeoff/landing (V/STOL) aircraft. The lift forces created by certain complicated flows are applied to V/STOL aircraft technologies. The jet induced force created by the JICF flow field can be applied to vehicle control, such as orbital flight maneuvering and re-entry of the Space Shuttle or potentially for high velocity and/or high altitude atmospheric vehicles.

One area of interest results from a transverse jet being directed into a hypersonic, laminar crossflow. The ensuing interaction of these two flows causes a change in surface pressure in the vicinity of the injector. When this modified pressure distribution is integrated a force is produced which can be several times larger than the nominal jet thrust. An application of this jet induced force is vehicle attitude control. When encountering situations where conventional aerodynamic surfaces cannot function properly, the system (of a jet induced force) is particularly advantageous. It is widely believed that this maybe due to the considerable aerodynamic heating effects associated with large flight speeds or possibly due to the low density of the surrounding medium. Also Vehicle Attitude Control is still possible from the thrust of a jet induced force, even when the external flow is so rarefied that the interaction force is negligible. Orbital Flight Maneuvering and re-entry of the Space Shuttle are well known examples of the control jet application, but other potential uses of this flow field can be applied to the high velocity and/or high altitude atmospheric vehicles [1].

The 72nd AGARD Fluid Dynamics Panel Meeting and Symposium on Computational and Experimental Assessment of Jets in Cross Flow [2] was the first meeting since 1981 where the primary theme was Jets in Cross Flow. The conference concluded with the general understanding that there is a need for improved prediction methods for the JICF problem. Investigations into transient flow features of a supersonic jet in a low speed cross flow [3], the separated flow generated by a vectored jet in a crossflow [4], and scalar mixing in the subsonic

jet in cross-flow [5] further define characteristics of the JICF flow field. Despite these efforts and others there are few results available for laminar hypersonic flows, a combination which will be encountered by re-entry and high altitude vehicles over some portion of their flight path. The experimental problem being simulated results from a series of nominally two dimensional experiments[1 ,6]. Interaction force data was obtained in a laminar, hypersonic freestream flow. To gain more knowledge about the influence of the jet within this part of the flow field, detailed measurements of the separated region have been made with various injectant species. Heat transfer and oil flow visualizations have also been used to elucidate the separation and reattachment process upstream and downstream of the jet.

Analysis

The experimental problem being modeled examines the interaction between a two-dimensional , Normal Sonic Jet of Nitrogen (or Methane) and a two-dimensional, Hypersonic Cross-Flow of Nitrogen over a flat plate. Performed in the Southampton University Light Piston Isentropic Compression Tube, a free stream of nitrogen at Mach 6.69 flows over an Isothermal Flat Plate with a sharp leading edge. This freestream interacts with a gas injected at room temperature via a normal slot jet located 0.0745 m from the leading edge of the plate (Figure 1).

The experimental problem was modeled and analyzed by the General Aerodynamic Simulation Program (GASP), version 2.2. The code GASP solves the fill Reynolds-averaged, compressible form of the Navier-Stokes, energy and species conservation equations. The code can be run in explicit, or implicit, space marching or elliptic modes. The governing equations are discretized using a finite volume approach and can be solved for one, two, and three dimensional models. The code can also utilize several thermodynamic, turbulent, and chemistry models [7]. In GASP the experimental problem was modeled using input decks which contain all of the relevant fluid flow information. The problem simulation and solution was performed on a grid which was generated by a FORTRAN program created outside of GASP. The numerical model consists of a two-dimensional, viscous, fully laminar flow with three different regions in the solution process: a upstream region, that is space marched; an injection region, that is globally iterated; and a downstream region, this is space marched. It is necessary to setup the solution process in this way because the injection into the flow will cause re-circulation of the flow in an area (injection region) surrounding the jet. This type of flow can only be captured by using a global iteration process. The flow upstream and downstream of this injection and or region can be simulated using a space marching solution scheme.

The inlet of the flow is supersonic, and therefore, the velocities, static temperature, static pressure, and species concentration are fixed values calculated from the conditions specified in [1]. The plate has the no-slip condition applied to it and is held at a constant temperature of 300K. This condition is appropriate for the high velocity and small test time encountered by the flow. The other boundaries of this two-dimensional control volume are setup to be extrapolated from the interior to first order accuracy. For the inviscid fluxes, the van Leer flux vector splitting with first order spatial accuracy is used. This inviscid flux calculation is used in conjunction with the MIN-MOD limiting algorithm. Once this more robust set-up creates an initial solution, the problem solution is completed by using Roe's flux difference splitting with Harten correction calculation with third order spatial accuracy. And the MIN-MOD limiting algorithm is used once again to control strong oscillations. The thermodynamics and chemistry options were set up to define the flow as an ideal, non-reacting, mixture of two species.

The computational simulation is performed with a two-dimensional grid. There are 201 grid points in the x-direction and 130 in the y-direction. An exponential stretching formula was

used to create appropriate clustering in the leading-edge region, the region downstream of injection and at the plate surface to capture all flow characteristics. And in the injection region the grid is uniformly spaced with very small increments to capture large flow variations. The grid was validated by comparing the Blasius solution curve from results of the present numerical model with this grid to the exact solution in [8].

Results and Discussions

All of the numerical simulations were performed on a Cray T-90 platform. In the results reported here, only injection of nitrogen is considered.

The main focus of this research is to produce a valid numerical model of the experimental problem of a normal sonic jet injected into a hypersonic cross-flow. The process of modeling the experimental problem was divided into two main parts. Part one consists of the verification and validation of a non-injection case of the problem, which is a simpler form of the problem. And in part two, the verification and validation of the more complicated injection case is considered.

The non-injection case, which is a modified version of the experimental problem where the normal sonic jet is turned off, was set-up and analyzed first. The non-injection case was set up to be space marched across the entire plate in the direction of the flow (x-direction.) This case can be space marched because it will not experience any of there-circulation of flow and other characteristics associated with injection into the freestream. The inviscid flux in the x-direction was set to a full flux with no splitting with a fully upwind second order accurate spatial accuracy. And in the y-direction Roe's flux difference splitting with Harten correction was used with a upwind-biased third order accurate setting. A MIN-MOD limiting was used. And only thin-layer contributions in the y-direction were considered.

The results showed that the non-injection simulation of the problem was very accurate. This model was verified using Van Driest calculations of the laminar compressible boundary layer on an flat plate for velocity and temperature profiles. The velocity and temperature profile behavior was very similar to that of the Van Driest solutions [8]. The model was also validated by comparing the measured surface pressure data [1] with the numerical model (figure 2). This comparison showed that the model predictions were accurate with only a $\pm$ **2°/0** error from the experimental values. Once the non-injection case was verified and validated to ensure that proper and accurate results were being produced, it was used to develop the injection case model.

The results from the non-injection case was used as an initial solution for simulating the injection case. The injection of the gas into the freestream results in re-circulation of the flow, therefore, the injection case is treated as an elliptic problem where the global iteration is appropriate. The model was also set up to consider the flow as a fully laminar, compressible flow with all the viscous terms in the x- and y-directions. The assumption of a laminar flow in our simulations is based on the observation of Ball [9] that the Reynolds number at the separation location was lower than its critical value and that there was a clear lack of evidence of transition in the thermographs record with no-injection.

The validation of the injection case model is currently underway. This case is a more complicated problem to solve than the non-injection case, so mesh sequencing was incorporated to facilitate the solution. Mesh sequencing is the process of converging a solution using a series of progressively finer meshes or grids. Vector plots of the present calculations show that re-circulation of the flow to be occurring. The injection case model is also being validated by comparing surface pressure values from the experiment to the surface pressure results from the numerical model (figure 3.) The model shows reasonably good agreement with the

measurements in the region between the leading edge and lip of injector. No measurements were reported for the region downstream of the injector.

Concluding Remarks

The non-injection case of the model shows very good agreement with the experimental measurements and in the injection case of the model good agreement with measurements particularly in the upstream region is observed. Agreement with the measured surface pressure may be improved with a more appropriate model setup. A multi-zone setup will be incorporated to model solve the injection case. This setup will have three zones, an injection region which will be globally iterated and two space marching region upstream and downstream of the injection region. Such a setup would allow for fine-tuning of the various parameters in the numerical model and the procedure is expected to yield better results.

Acknowledgments

The numerical computations reported here were performed on Cray T90 systems and were made possible by the North Carolina Supercomputing Center in Research Triangle Park, NC. Financial support provided by the grant Nagw-2924 from NASA in gratefully acknowledged.

References

[1] Powrie, H.E.G., Ball, G.J., East, R.A. "Comparison of the Interactions of Two and Three Dimensional Transverse Jets with a Hypersonic Free Stream", AGARD Conference Proceedings 534, November, 1993, pp. 20-1--20-8.

[2] AGARD Conference Proceedings 534, Computational and Experimental Assessment of Jets in Cross Flow, North Atlantic Treaty Organization, November 1993.

[3] X. Zhang, D.W. Hurst, and G.M. Lilley, "Transient Flow Features of a Supersonic Jet in a Low Speed Cross Flow", AGARD Conference Proceedings 534, Reference 3, North Atlantic Treaty Organization, November 1993.

[4]A. Krothapalli and C. Shih, "Separated Flow Generated by a Vectored Jet in a Crossflow", AGARD Conference Proceedings 534, Reference 5, North Atlantic Treaty Organization, November 1993.

[5] S.H. Smith, A. Lozano, M.G. Mungal, and R. K. Hanson, "Scalar Mixing in the Subsonic Jet in Crossflow", Reference 6, AGARD Conference Proceedings 534, Reference 5, North Atlantic Treaty Organization, November 1993.

[6] Powrie, H.E.G, Ball, G.J., and East, R.A., "Experimental Study of a Two Dimensional Control Jet with a Hypersonic Flow", IUTAM Conference on Aerothermochemistry of Spacecraft and Associated Hypersonic Flows, Marseille, September 1992

[7] McGrory, William D., Slack, David C., Applebaum, Michael P., Walters, Robert W., Gasp version 2.2 User's Manual, AeroSoft, Inc., 1993.

[8] Anderson, John D., Jr., Computational Fluid Dynamics, McGraw-Hill, Inc., New York, 1995.

[9] Ball, G.J., personal correspondence, August 5,1996.

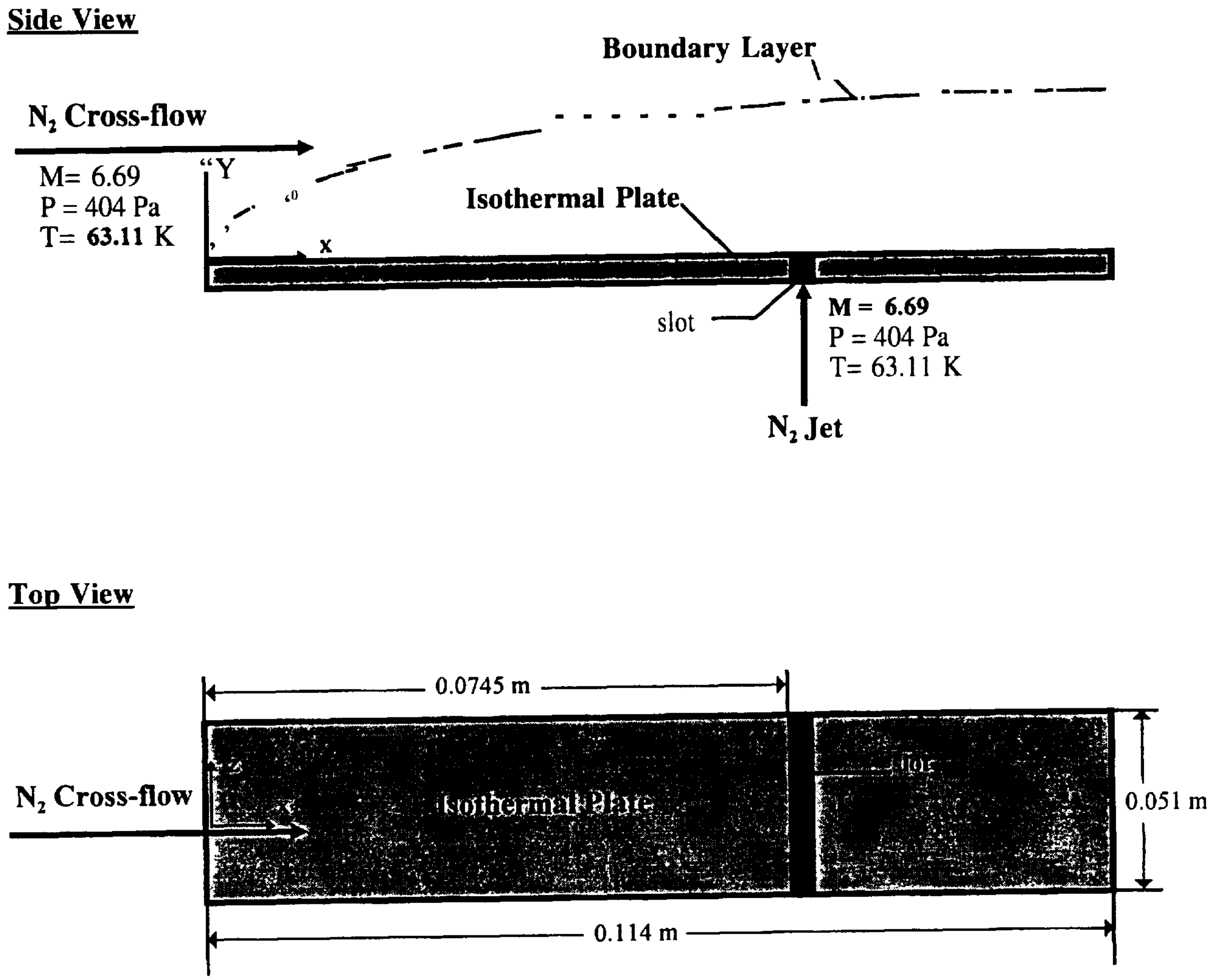


Figure 1. Schematic of experimental problem to be simulated numerically.

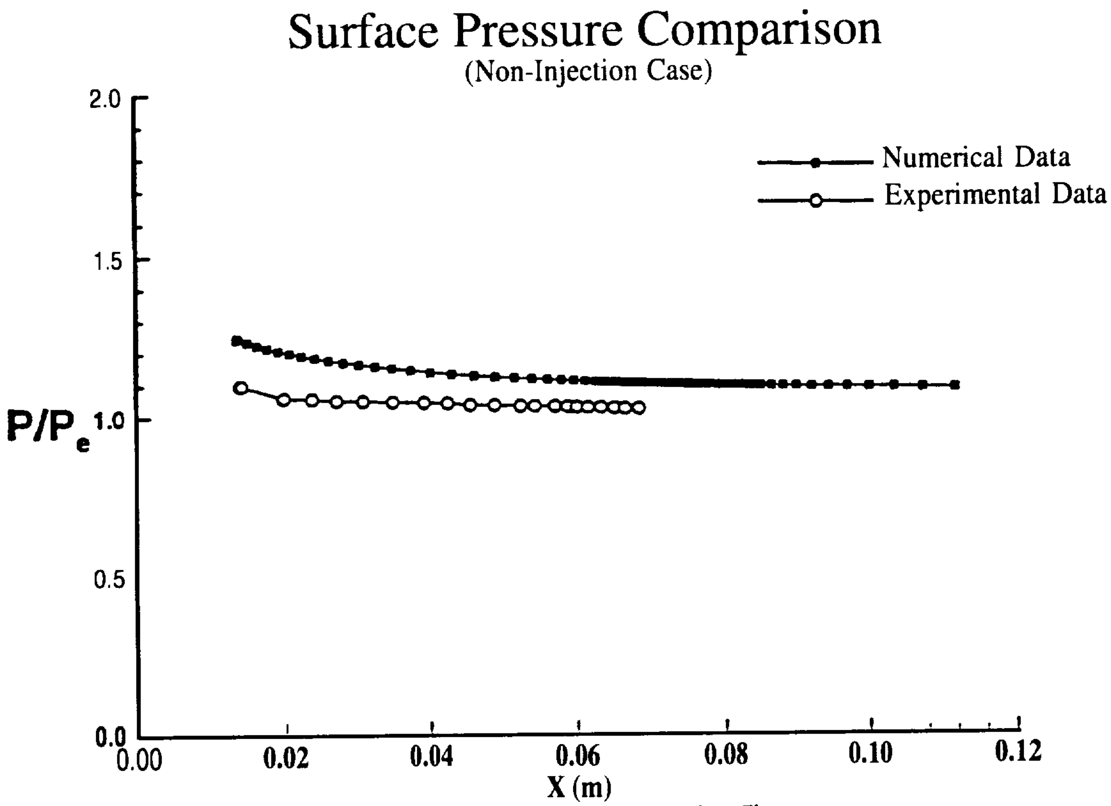


Figure 2. Surface Pressure Comparison Plot

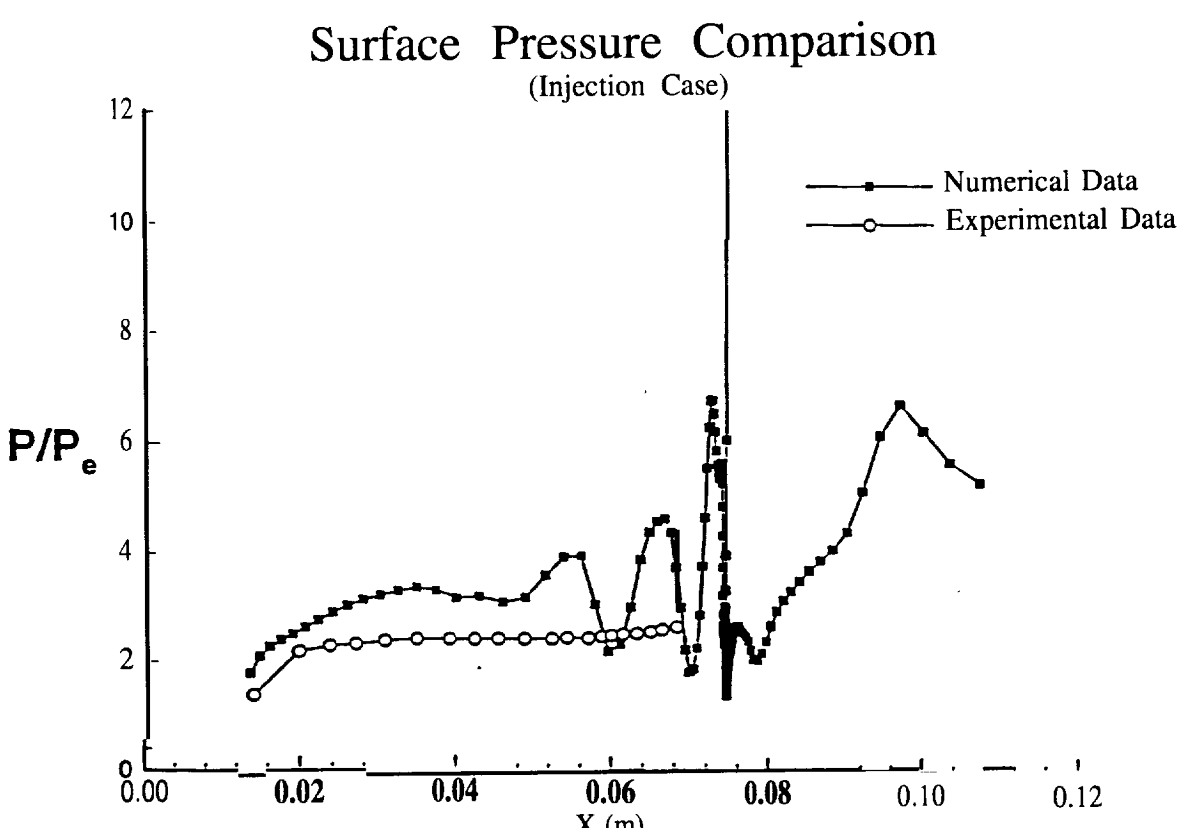


Figure 3. Surface Pressure Comparison Plot (Injection)

URC97063

Supervised Classification Techniques for Hyperspectral Data*

Luis O. Jimenez
Tropical Center for Earth and Space Studies
Electrical & Computer Engineering Department
University of Puerto Rico, Mayaguez, Puerto Rico
jimenez@exodo.upr.clu.edu

N/ASA
5 3/82

Abstract

The recent development of more sophisticated remote sensing systems enables the measurement of radiation in many mm-e spectral intervals than previous possible. An example of this technology is the AVIRIS system, which collects image data in 220 bands. The increased dimensionality of such hyperspectral data provides a challenge to the current techniques for analyzing such data. Human experience in three dimensional space tends to mislead one's intuition of geometrical and statistical properties in high dimensional space, properties which must guide our choices in the data analysis process. In this paper high dimensional space properties are mentioned with their implication for high dimensional data analysis in order to illuminate the next steps that need to be taken for the next generation of hyperspectral data classifiers.

I. Introduction

The complexity of dimensionality has been known for more than three decades, and its impact varies from one field to another. In combinatorial optimization over many dimensions, it is seen as an exponential growth of the computational effort with the number of dimensions. In statistics, it manifests itself as a problem with parameter or density estimation due to the paucity of data. The negative effect of this paucity results from some geometrical, statistical and asymptotical properties of high dimensional feature space. These characteristics exhibit surprising behavior of data in higher dimensions.

There are many assumptions that we make about characteristics of lower dimensional spaces based on our experience in three dimensional Euclidean space. There is a conceptual barrier that makes it difficult to have proper intuition of the properties of high dimensional space and its consequences in high dimensional data behavior. Most of the assumptions that are important for statistical purposes we tend to relate to our three dimensional space intuition, for example, as to where the concentration of volume is of such figures as cubes, spheres, and ellipsoids or where the data concentration is in known density function families such as normal and uniform. Other important perceptions that are relevant for statistical analysis are, for example, how the diagonals relate to the coordinates, the number of labeled samples required for supervised classification, the assumption of normality in data, and the importance of mean and covariance difference in the process of discrimination among different statistical classes. In the next section some characteristics of high dimensional space will be mentioned, and their impact in supervised classification data analysis will be discussed. Most of these properties do not fit our experience in three dimensional Euclidean space as mentioned before.

II. Geometrical, Statistical and Asymptotical Properties

In this section we illustrate some unusual or unexpected hyperspace characteristics including a discussion of its implications for supervised classification. These illustrations are intended to show that higher dimensional space is quite different from the three dimensional space with which we are familiar.

As dimensionality increases:

* Work reported herein was funded in part by NASA Grant NAGW-3924.

A. The volume of a hypercube concentrates in the corners and the volume of a hypersphere concentrates in an outside shell [Scott 1992].

These characteristics have two important consequences for high dimensional data that appear immediately. The first one is that high dimensional space is mostly empty, which implies that multivariate data in R^d is usually in a lower dimensional structure. As a consequence high dimensional data can be projected to a lower dimensional subspace without losing significant information in terms of separability among the different statistical classes. The second consequence of the foregoing, is that normally distributed data will have a tendency to concentrate in the tails; similarly, uniformly distributed data will be more likely to be collected in the corners, making density estimation more difficult. Local neighborhoods are almost surely empty, requiring the bandwidth of estimation to be large and producing the effect of losing detailed density estimation. Support for this tendency can be found in the statistical behavior of normally and uniformly distributed multivariate data at high dimensionality. It is expected that as the dimensionality increases the data will concentrate in an outside shell. As the number of dimensions increases that shell will increase its distance from the origin as well. Under these circumstances it would be difficult to implement any density estimation procedure and to obtain accurate results. Generally nonparametric approaches will have even greater problems with high dimensional data.

B. The required number of labeled samples for supervised classification increases as a function of dimensionality.

Fukunaga [Fukunaga 1989] proves that the required number of training samples is linearly related to the dimensionality for a linear classifier and to the square of the dimensionality for a quadratic classifier. That fact is very relevant, especially since experiments have demonstrated that there are circumstances where second order statistics are more relevant than first order statistics in discriminating among classes in high dimensional data [Lee and Landgrebe, July 1993], In terms of nonparametric classifiers the situation is even more severe. It has been estimated that as the number of dimensions increases, the sample size needs to increase exponentially in order to have an effective estimate of multivariate densities [Scott 1992, pp 208-212] [Hwang, Lay, Lippman 1994],

It is to be expected that high dimensional data contains more information. At the same time the above characteristics tell us that it is difficult with the current techniques, which are usually based on computations at fill dimensionality, to extract such information unless the available labeled data is substantial. A concrete example of this is the so-called Hughes phenomena. Hughes proved that with a limited number of training samples there is a penalty in classification accuracy as the number of features increases beyond some point [Hughes 1968].

C. For most high dimensional data sets', low linear projections have the tendency to be normal, or a combination of normal distributions, as the dimension increases.

That is a significant characteristic of high dimensional data that is quite relevant to its analysis. It has been proved [Diaconis and Freedman 1984] [Hall and Li 1993] that as the dimensionality tends to infinity, lower dimensional linear projections will approach a normality model with probability approaching one (see Figure 6). Normality in this case implies a normal or a combination of normal distributions.

In all the cases above we can see the advantage of developing an algorithm that will estimate the projection directions that separate the explicitly defined classes, doing the computations in a lower dimensional space. The vectors that it computes will separate the classes, and at the same time, the explicitly defined classes will behave asymptotically more like a normal distribution. The assumption of normality will be better grounded in the projected subspace than in fill dimensionality.

D. The role of the second order statistics become as important as the first order statistics.

Lee and Landgrebe [Lee and Landgrebe July 1993] performed an experiment where they classified some high dimensional data in order to see the relative role that first and second order statistics played.

In that particular experiment as the number of dimension grew the role played by the second order statistics increased in discriminating among classes. Under these circumstances, the shape of the distribution given by the second order statistics becomes as important as the location provided by the first order statistics.

III. High dimensional characteristics implications for supervised classification

Based on the characteristics of high dimensional data that the volume of hypercubes have a tendency to concentrates in the corners, and in a hyperellipsoid in an outside shell, it is apparent that high dimensional space is mostly empty, and multivariate data is usually in a lower dimensional structure. As a consequence it is possible to reduce the dimensionality without losing significant information and separability. Due to the difficulties of density estimation in nonparametric approaches, a parametric version of data analysis algorithms maybe expected to provide better performance where only limited numbers of labeled samples are available to provide the needed a priori information.

The increased number of labeled samples required for supervised classification as the dimensionality increases presents a problem to current feature extraction algorithms where computation is done at fill dimensionality, e.g. Principal Components, Discriminant Analysis and Decision Boundary Feature Extraction [Lee & Landgrebe, April 1993]. A new method is required that, instead of doing the computation at full dimensionality, computes in a lower dimensional subspace. Performing the computation in a lower dimensional subspace that is a result of a linear projection from the original high dimensional space will make the assumption of normality better grounded in reality, giving a better parameter estimation, and better classification accuracy.

A preprocessing method of high dimensional data based on such characteristics has been developed based on a technique called Projection Pursuit. The preprocessing method is called Parametric Projection Pursuit [Jimenez and Landgrebe IGARSS 95] [Jimenez and Landgrebe SMC 95].

Parametric Projection Pursuit reduces the dimensionality of the data maintaining as much information as possible by optimizing a Projection Index that is a measure of separability. The projection index that is used is the minimum Bhattacharyya distance among the classes, taking in consideration first and second order characteristics. The calculation is performed in the lower dimensional subspace where the data is to be projected. Such preprocessing is used before a feature extraction algorithm and classification process, as shown in Figure 1.

In Figure 1 the different feature spaces have been named with Greek letters in order to avoid confusion. Φ is the original high dimensional space. Γ is the subspace resulting from a class-conditional linear projection from Φ using a preprocessing algorithm, e.g. Parametric Projection Pursuit. Y is the result of a feature extraction method. Y could be projected directly from Φ or, if preprocessing is used, it is projected from 17. Finally Ω is a one dimensional space that is a result of classification of data from Y space. Note that the three procedures, preprocessing, feature extraction and classification use labeled samples as a priori information.

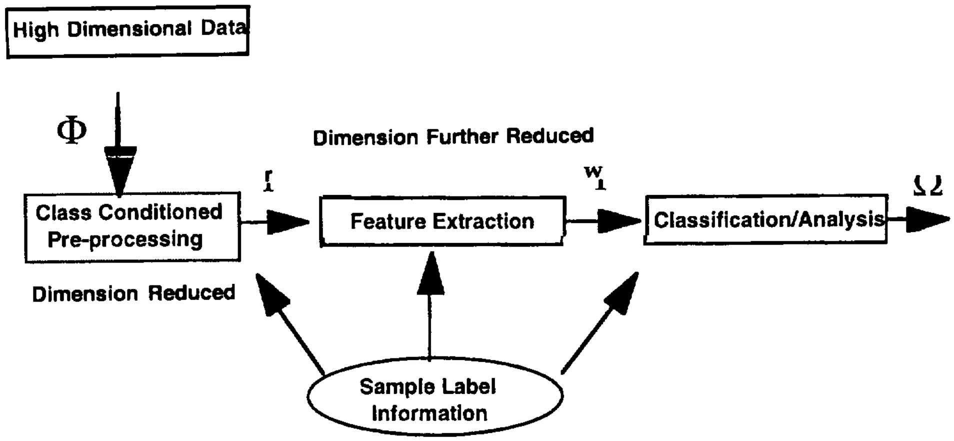


Figure 1. Classification of high dimensional data including preprocessing

IV. Experiment

In order to see the relevance of high dimensional geometrical and statistical properties for high dimensional data analysis purposes two experiments were designed. In both experiments a comparison is provided between high dimensional feature extraction and the method that uses a Parametric Projection Pursuit based preprocessing to reduce the dimensionality before a feature extraction method is used. The multispectral data used in these experiments are a segment of AVIRIS data taken of NW Indiana's Indian Pine test site. From the original 220 spectral channels 200 were used, discarding the atmospheric absorption bands.

The classification task for several classes in this and the next experiment are particularly difficult ones. The data were collected early in the growing season when the canopy of both corn and soybeans covered only about 5% of the area, There were three levels of tillage, no till in which there would be a great deal of residue on the soil surface from last year's crop, minimum till leaving a moderate amount of residue, and clean till for which there would be little or no residue. Add to this the normal amount of spectral variability due to the varying soil types present in the fields. Thus the 95% background would be highly variable, as compared to the relatively small difference in spectral response between corn and soybeans.

In this experiment four classes were defined: corn, corn-notill, soybean-rein, soybean-notill. The total number of training samples is 179 (less than the number of bands used) and the total number of test samples is 3501. Observe that this is an extreme case that is used to show the potentials of Parametric Projection Pursuit. Two types of dimensional reduction algorithms were used. The first is Discriminant Analysis (DA 200-3) that reduces the dimensionality from 200 to 3. It directly projects the data from Φ space to Y subspace. In the second method Parametric Projection Pursuit was used to reduce the dimensionality from 200 to 22. It projected the data from the Φ space to the Γ subspace. After that preprocessing method was used, Discriminant Analysis was used (PPDA 200-3) in order to linearly project the data from the Γ subspace to the Ψ subspace. As mentioned before, this has the advantage of doing the computation with the same number of training samples but at lower dimensionality. In both cases the best three features were used for classification purposes.

Four types of classifiers were used. The first one is ML classifier, the second is ML with 2% threshold. The third classifier is a spectral-spatial classifier named ECHO [Kettig & Landgrebe 1976] [Landgrebe 1980] and the fourth is ECHO with a 2% threshold. In the second and the fourth, a threshold was applied to the standard classifiers whereby in case of true normal

distributions of the data, 2% of the least likely points will be thresholded. These 2% thresholds provide one indication of how well the data fit the normal model.

The results are shown in Figure 2. Parametric Projection Pursuit followed by Discriminant Analysis at lower dimensionality performed substantially better than using Discriminant Analysis at fill dimensionality. The application of a threshold to Discriminant Analysis at full dimensionality reduced its classification accuracy more severely than when a threshold was applied in the case where Projection Pursuit was first applied, followed by Discriminant Analysis at lower dimensionality. This is due to Parametric Projection Pursuit preprocessing being better fitted to the assumption of normality.

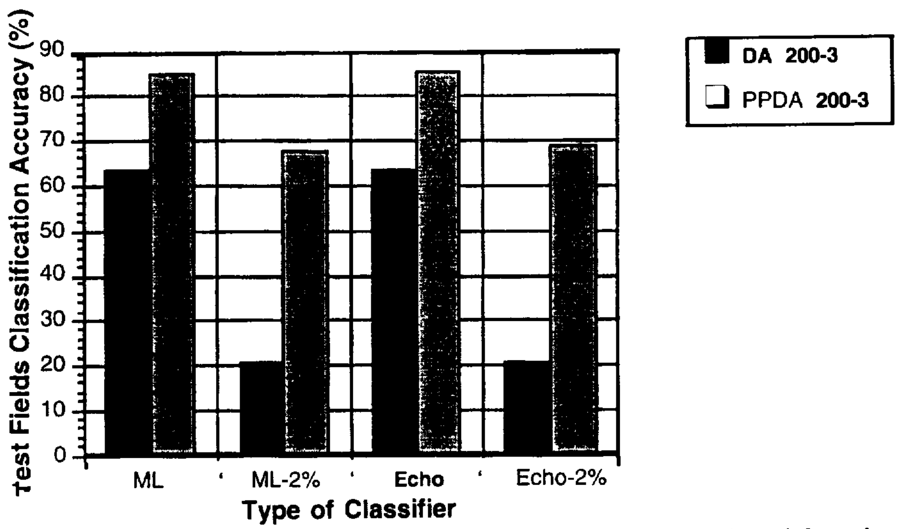


Figure 2. Test fields classification accuracy for two feature extraction methods and four classifiers.

Observe how significantly the performance of classifiers with 2% thresholds improves when using Parametric Projection Pursuit. The reason is that making the computation at low dimensional space, Γ, the assumption of normality has greater validity. In the case of having less samples and classes Discriminant Analysis will be significantly affected by the high dimensional geometrical and statistical characteristics. The next experiment will show this difficulty.

VI. Conclusion

In this section we will consider some implications of what has been discussed for supervised classification. In terms of parameter estimation, a large number of samples are required to make a given estimation in multispectral data to adequate precision. In a nonparametric approach, the number of samples required to satisfactorily estimate the density is even greater. Both kinds of estimations confront the problem of high dimensional space characteristics. As a consequence, it is desirable to project the data to a lower dimensional space where high-dimensional geometric characteristics and the Hughes phenomena are reduced. Commonly used techniques such as Principal Components, Discriminant Analysis, and Decision Boundary Feature Extraction have the disadvantage of requiring computations at full dimensionality in which the required number of labeled samples is very large. The procedures use estimated statistics that are not necessarily accurate. Another problem is the assumption of normality. Nothing guarantees that at full dimensionality, that model fits well.

It has been shown that high dimensional spaces are mostly empty, indicating that the data structures involved exist primarily in a subspace. The problem is which subspace it is to be

found in is situation-specific. Thus the goal is to reduce the dimensionality of the data to the right subspace without losing separability information. The approach is to make the computations in a lower dimensional space, i.e. in Γ instead of Φ, where the projected data produce a maximally separable structure and which, in turn, avoids the problem of dimensionality in the face of the limited number of training samples. Further, a linear projection to a lower dimensional subspace will make the assumption of normality in the Γ subspace more suitable than in the original Φ. In such a lower dimensional subspace any method used for feature extraction could be used before a final classification of data, even those that have the assumption of normality.

References

Chulhee Lee and David A. Landgrebe, "Feature Extraction Based On Decision Boundaries," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol 15, No. 4, April 1993, p p 388-400.

Chulhee Lee and David A. Landgrebe, "Analyzing High Dimensional Multispectral Data," *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 31, No, 4, pp 792-800, July, 1993.

Diaconis, P. , Freedman, D. "Asymptotic of Graphical Projection Pursuit." The Annals of Statistics Vol 12, No 3 (1984): pp 793-815.

Fukunaga, K. "Introduction to Statistical Pattern Recognition," San Diego, California, Academic Press, Inc., 1990.

Hall, P., Li, K. "On Almost Linearity Of Low Dimensional Projections From High Dimensional Data." The Annals of Statistics, Vol. 21, No. 2 (1993): pp 867-889.

Hughes, G. F., "On the mean accuracy of statistical pattern recognizes," IEEE Transactions on Information Theory, Vol. IT-14, No. 1, January 1968.

Hwang, J., Lay, S., Lippman, A., "Nonparametric Multivariate Density Estimation: A Comparative Study.", IEEE Transactions on Signal Processing, Vol. 42, No. 10, 1994, p p 2795-2810.

Jimenez, L., Landgrebe, D., "Projection Pursuit For High Dimensional Feature Reduction: Parallel And Sequential Approaches," presented at the International Geoscience and Remote Sensing Symposium (IGARSS'95), Florence Italy, July 10-14, 1995.

Jimenez, L., Landgrebe, D., "Projection Pursuit in High Dimensional Data Reduction: Initial Conditions, Feature Selection and the Assumption of Normality", To be presented at IEEE International Conference on Systems, Man and Cybernetics (SMC 95), Vancouver Canada, October 22-25, 1995.

R. L. Kettig and D. A. Landgrebe, "Computer Classification of Remotely Sensed Multispectral Image Data by Extraction and Classification of Homogeneous Objects," *IEEE Transactions on Geoscience Electronics,* Volume GE-14, No. 1, pp. 19-26, January 1976.

D.A. Landgrebe, "The Development of a Spectral-Spatial Classifier for Earth Observational Data," *Pattern Recognition,* Vol. 12, No. 3, pp. 165-175,1980.

Scott, D. W. "Multivariate Density Estimation." New York: John Wiley& Sons, 1992.

Turbulent Distortion of Condensate Accretion

R. Hazoumé,*, J. Chabi Orou,** J. A. Johnson III
CeNNAs, Florida A&M University, Tallahassee, FL 32310 (USA)

Abstract

When a simple model for the relationship between the density-temperature fluctuation correlation and mean values is used, we determine that the rate of change of turbulent intensity can influence directly the accretion rate of droplets.

Considerable interest exists in the accretion rate for condensates in nonequilibrium flow with icing[1] and the potential role which reactant accretion can play in nonequilibrium exothermic reactant processes.[2] Turbulence is thought to play an important role in such flows. It has already been experimentally determined that turbulence influences the sizes of droplets in the heterogeneous nucleation of supersaturated vapors.[3] This paper addresses the issue of the possible influence of turbulence on the accretion rate of droplets.

According to a nonequilibrium model developed earlier[4] the droplet growth rate for flow with changing local pressure is given by:

$$\frac{dr(t)}{dt} = -3BN_n\frac{d}{dt}\log\frac{P(t)}{P_\infty} \tag{1}$$

with B= (2s)/[r_cT(R/m_n)] where s = surface tension coefficient r_c = condensate density; and m_n = molecular weight of vapor. In this approach, we use the hardcore model for a liquid droplet in which dN, the number of simple molecules of water in the shell between r and dr, is given by $dN = (k/r^n)(4\pi r^2)dr$. N_n is a correction coefficient such that $0 \le N_n \le 4$ and $N_n = [(n-3)^{n-3}]/[(n-4)^{n-2}]$ for n>4. P_∞ is obtained through the Clausius-Clapeyron equation. For adiabatic flow of a compressible ideal gas the mean value of the square of the local speed of sound is defined as

$$\overline{b^2} = \left(\frac{\partial\overline{P}}{\partial\overline{\rho}}\right)_{\overline{s},\overline{g}} = \gamma\frac{\overline{P}}{\overline{\rho}} \tag{2}$$

where: g = C_p/C_v is the ratio of the specific heats; $\overline{g}$ is the mean value of the condensate mass fraction of water; $\overline{P}$ is the mean pressure function; $\overline{s}$ is the mean entropy; and $\overline{\rho}$ is the mean density function.

We can readily write from (2) that

$$\frac{\overline{P}}{\overline{\rho}^{\gamma}} = cnst \quad \text{and} \quad \frac{\overline{P}}{(\overline{b^2})^{\frac{\gamma}{(\gamma-1)}}} = cnst \tag{3}$$

Equations (2) and (3) were written assuming there are fluctuations of the principle physical properties around their mean values as, e.g., g = $\overline{g}$ + g' with $\overline{g'} = O$. However $\overline{g'^2}$ #O. This means that we can consider $P = \overline{P} + P'$, $\rho = \overline{\rho} + \rho'$, g = $\overline{g} + g'$, and $T = \overline{T} + T'$. The equation of state

$$P = \rho\left(\frac{1-\omega_o}{\mu_i} + \frac{\omega_o - g}{\mu_v}\right)RT \tag{4}$$

becomes

$$\overline{P} + P' = (\overline{\rho} + \rho')\left(\frac{1-\omega_o}{\mu_i} + \frac{\omega_o - \overline{g} - g'}{\mu_v}\right)R(\overline{T} + T') \tag{5}$$

with ω_o=the initial humidity, μ_i=the molecular mass of the carrier gas, and R=the ideal gas constant. Averaging equation (4) yields:

$$\overline{P} = \left(\frac{1-\omega_o}{\mu_i} + \frac{\omega_o\ \overline{g}}{\mu_v}\right)R(\overline{T}\overline{\rho} + \overline{\#T'}) \tag{6}$$

where a new function, $\overline{\rho' T'}$, has appeared, this is the density-temperature correlation function involving the fluctuations of density and temperature. We can use this result in eq. (5) above.

Now we make the important simple assumption

$$\overline{\rho'T'} = -\lambda\overline{\rho}\overline{T}. \tag{7}$$

in which eq. (7) is a definition of λ; with this definition, λ is without consequences for any specific experiment. The equation of state then becomes

$$\frac{\overline{P}}{\overline{\rho}} = l(\overline{g})(1-\lambda)R\overline{T} = \frac{b^2}{\gamma} \tag{8}$$

where

$$l(\overline{g}) = \frac{1-\omega_o}{\mu_i} + \frac{\omega_o - \overline{g}}{\mu_v}. \tag{9}$$

We must also make it clear that the physical properties of interest are functions of space and time. That is,

$$\overline{T} = \overline{T}(x_i,t), \overline{P} = \overline{P}(x_i,t), \overline{\rho} = \overline{\rho}(x_i,t), \overline{g} = \overline{g}(x_i,t), \lambda = \lambda(x_i,t) \tag{lo}$$

where i = 1,2,3, xi is the i component of the space variable, and t is the time. To use equation (10), we can use as the simplest adequate physically reasonable model for $\lambda(x_i,t)$ to be $-\lambda(x_i, t\ ) = \exp - ft$. The exponential character of the decay is consistent with the usual long-term behavior of turbulent correlations. With this model, the density temperature correlation function depends only on time and obeys a decay law which is exponential. Finally, it is then usual (with U and L as velocity and length parameters respectively) to express all these relations in a non-dimensional form: vi $=u_i/U$; $y_i=x_i/L$; and t=tU/L.

We can now include the influence of fluctuations in temperature (as well as the implicit fluctuations in P and p) in a turbulence distorted refinement of the local mean square of the local speed of sound $\overline{a^2(y_2,\tau)}$ by explicitly requiring $\frac{\overline{P}(y_2,\tau)}{\overline{\rho}(y_2,\tau)} = \frac{\overline{a^2}}{\gamma}$. This form is reconciled with equation (8) above by

$$\frac{\overline{b^2}}{\gamma} = \frac{\overline{a^2}}{\gamma}(1-\lambda(y_2,\tau)) \tag{11}$$

We have shown in equation (3) that $\overline{P}\alpha(b^2)^{\frac{\gamma}{(\gamma-1)}}$. Thus

$$\frac{\partial}{\partial\tau}Log\overline{P}(y_i\ ,\ \tau) = \frac{\gamma}{(\gamma-1)}\left[\frac{\partial\overline{a^2(y_2,\tau)}}{\partial\tau} + \frac{\partial}{\partial\tau}Log(1-\lambda)\right] \tag{12}$$

$$= \frac{\gamma}{\gamma-1}\{\frac{\partial}{\partial\tau}Log\overline{a^2(y_2,\tau)} - \frac{f}{\exp f\tau + 1}\}$$

in which f is the real positive constant introduced in the definition of 1. We finally obtain from equation (1)

$$\frac{dr(\tau)}{d\tau} = \frac{6\sigma \quad (n \quad - \quad 3)^{n-3} \quad \gamma \quad f}{\rho_c(\frac{R}{\mu_v})\overline{T}\ (n\ 4)^{n-2}\ (\gamma\ -\ 1)\exp f\tau\ +\ 1} - \frac{\partial Log\overline{a^2(y_2,\tau)}}{\partial\tau}\ \} \tag{13}$$

When we define the instantaneous turbulent intensity as proportional to $\overline{a^2(y_2,\tau)}$, then the interpretation of equation (13) is straightforward. (i)At small values of t and constant turbulent intensity, ($\partial\overline{a^2(y_2,\tau)}/\partial\tau=0$), the rate of change in droplet size is determined entirely by the density correlation exponential factor f. (ii)At large values oft and constant turbulent intensity, ($\partial\overline{a^2(y_2,\tau)}/\partial\tau=0$), the rate of change in droplet size has no dependence on turbulence at all. (iii)If the turbulent intensity is not constant, i.e., ($\partial\overline{a^2(y_2,\tau)}/\partial\tau\ \#O$), then the rate of change in droplet size is sensitive to the turbulence at all times. The case $\partial\overline{a^2(\ y_2,\tau)}/\partial\tau \neq 0$ requires that a measure of the actual evolution in the strength of turbulent fluctuations be determined and that a model for determining the connection between the mean square local speed of sound and the evolution of turbulence be determined.

The overall features just described are summarized in Fig. 1. It has already been shown[2] that, at fried thermodynamic conditions, the size of a droplet in condensing flow is influenced by the strength of the local turbulence and, by inference, the value of the Reynolds. Equation (13) above sets conditions for the sensitivity of the condensate accretion to turbulence, based on basic physics principles, which can now be tested by experiments.

This research was supported in print by NASA Grants NAGW-2930 and NAG 2-291 and by DOE Contract DE-FG04-95AL87309 to Florida A&M University.

References

*Visiting Research Scientist. Permanent Address: Université Nationale du Benin, Cotonou, Benin
**Graduate Research Assistant. Current Address: Université Nationale du Benin, Cotonou, Benin

[1]T. Cebeci, AIAA Journal, 33, 1995, pp. 1351-2
[2]J. A. Johnson III, L. E. Johnson and X. Lu, Phys. Fluids A, 2, 2002 (1990)
[3]U.De Silva, A. Gardner, J. A. Johnson III, AIAA Journal, 33_ 1995, pp. 368-370.
[4]R. Hazoumé and J. Chabi Orou, "Pressure Sensitivity in the Droplet Growth Rate in Nonequilibrium Condensation: in Mathematical and Experimental Developments for Nonequilibrium Statistical Systems, Institut de Mathematiques et de Sciences Physiques, Université Nationale du Benin, May 25-27,1994, Contonu, Benin (to be published)

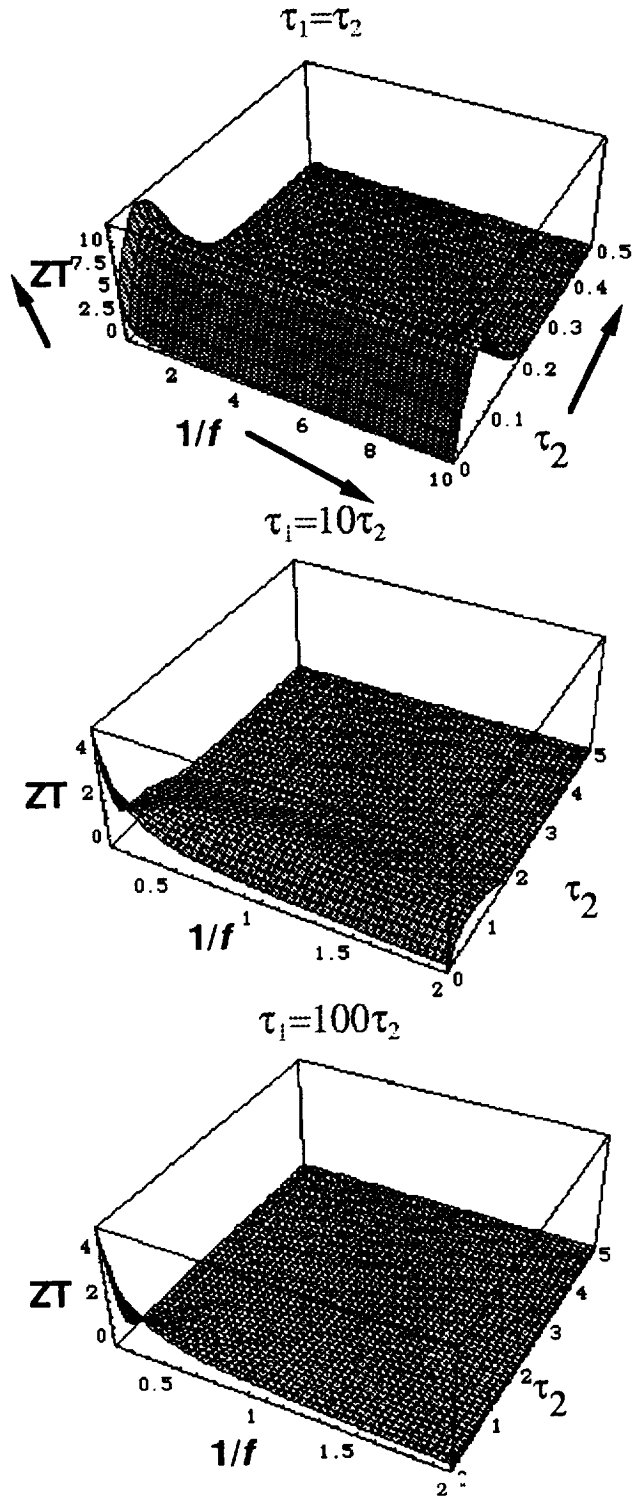


Figure 1. Sensitivity in the Evolution of Turbulent Distortion of Droplet Growth Rate with the Persistence of the Turbulence Pulse. The symbol f is defined in the text; τ_2 is the laboratory time; τ_1 is the persistence time of adistortion in the turbulence; and ZT is the product of the left hand side of eq. (13) with the prefactors of the right hand side's bracketed expression. The plots show the differences between a short puke, in the top display, and a long and relatively unchanging turbulence in the lower display.

INTELLIGENT PARAMETER ESTIMATION OF NASA/JPL FLEXIBLE BEAM DAMPING COEFFICIENTS: ARTIFICIAL NEURAL NETWORKS

MARK A. JOHNSON
ERIC HAMKE

NASA Center for Autonomous Control Engineering (ACE)
Department of Electrical and Computer Engineering
University of New Mexico, Albuquerque, NM 87131
Email: ace@pajarito.unm. edu

ABSTRACT

NASA/JPL Flexible Beam dumping coefficients are estimated by multilayer perception artificial neural networks (ANNs). Flexible beam damping coefficients are estimated from the difference between actual and modeled open loop beam behavior. Spatial, temporal, and spatio-temporal coefficient estimation is performed. Dynamically estimated dumping parameters enable the simulated flexible beam behavior to emulate actual beam behavior where analytically derived static coefficients fail. Results indicate the concept applicability and provide a basis for future work.

1. INTRODUCTION, MOTIVATION, AND APPROACH

Control of structures in space is a critical problem to resolve as it is applicable to anything planned for space stations (orbital or surface), satellites, and inter/intra planetary/stellar vessels. One key to controlling these structures is knowledge of system parameters as they undergo spatio-temporal changes. Parameter estimation is used to provide useful information on important, indirectly observable system parameters. One goal in parameter estimation is to reduce computational complexity while achieving acceptable performance. Intelligent methods provide estimates which are at least as good as conventional analytical methods while reducing computational complexity and estimation delay. Current core intelligent methods are artificial neural networks, fuzzy logic, genetic algorithms, genetic programming, expert systems, and probabilistic reasoning [1]. Performing parameter estimation with potential direct application to adaptive flexible space system(s) control using intelligent techniques is the focus of this paper.

Intelligent methods of estimating parameters are a key component to controlling any future flexible space system. However, little exists in literature concerning experimental applications of intelligent approaches to parameter estimation, which motivates the application of an intelligent method to estimate parameters and is a brief synopsis of a portion of the work contained in references 2 and 3. Section 2 begins with the problem statement, followed by the approach taken to solve the problem using ANNs being presented in section 3. Section 4 gives and discusses some experimental results after which conclusions, recommendations, and future directions are presented.

2. PROBLEM STATEMENT

The experimental test bed is graphically depicted in Figure 1. The problem statement is, as follows:

Given: 1) A "distributed-parameter, infinite-degree-of-freedom "plant""
(NASA/ACE laboratory, NASA/JPL Flexible Beam testbed) [4];
2) A fifty element model of the flexible beam;
3) Two "levitators" at elements 20 and 40 which counterbalance gravity effects and are coupled to optical encoders; and

and 40) during actuator excitation;

Estimate, using intelligent methods, the damping coefficients from actuator and sensor inputs,

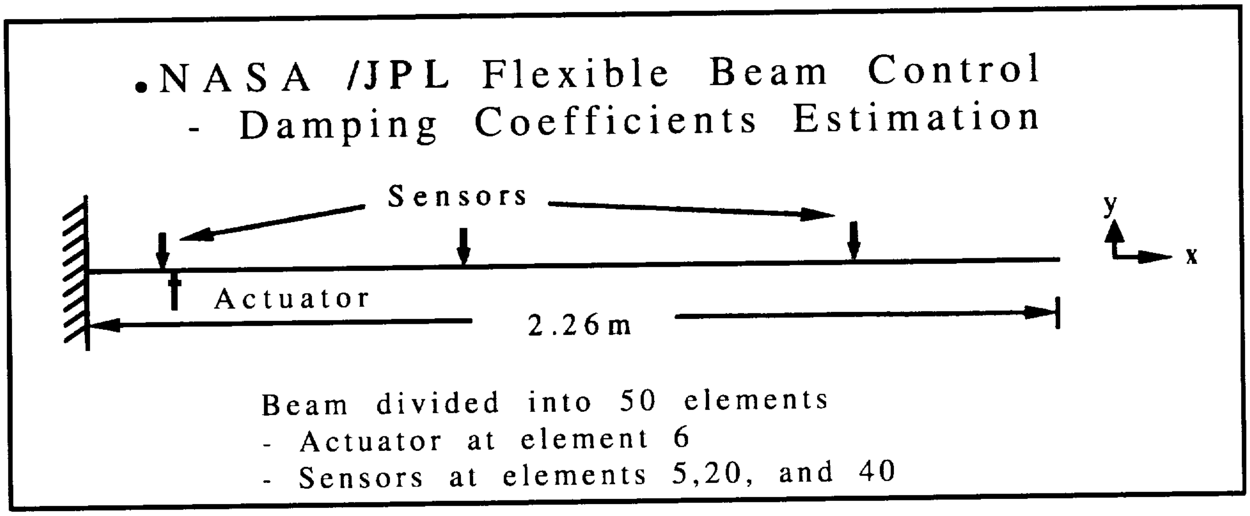


Figure 1. NASA/ACE Flexible Beam Testbed

3. INTELLIGENT PARAMETER ESTIMATION

The basic parameter estimation problem consists of five (5) components [5]. The five components are: 1) The variable(s) to be estimated; 2) Available measurements or observations; 3) Mathematical model describing how the measurements are related to the variable(s) being estimated; 4) Mathematical model of the uncertainties present; and 5) Performance evaluation criteria. Variable(s) to be estimated are the damping coefficients for elements 5, 20, and 40. Each is used due to having sensors on the beam at these location from which direct measurements are made. Only elements 20 and 40 are under direct suspension in an attempt to counteract the effects of gravity. Element 5 is next to element 6, the actuator. One of the advantages of using intelligent techniques is the nonparametric nature of the solution. In this sense the mathematical models of components 3 and 4 are not necessary (see [4] for derivation of mathematical model).

To understand the evaluation criteria for the problem one first needs to understand the solution approach. Damping coefficients and their variations within the Flexible Beam structure are not observable. Flexible Beam and model inputs/outputs are observable and available. The hypothesis is to make the model behave like the actual open loop Flexible Beam by dynamically adjusting the

damping coefficients. As the error between the model versus Beam behavior goes to zero, the damping coefficients should converge to their appropriate values since they are the only parameters changing within the model. Neural network weights are changed according the mean squared error of the model versus actual beam behavior after each epoch of training. The experimental setup for training the neural networks is shown in Figure 2.

The Flexible Beam analytical model is a linearization of the actual nonlinear beam (a cantilever beam pinned at one end) and all unmodeled dynamics are assumed to be included within the damping coefficients. The fourth order differential equations describing the beam model are found in any mechanical engineering text (see [12]). The equations are modified around those elements with levitators to account for gravity losses, actuators and sensors are assumed to present negligible mass. The model is realized in MATLAB and SIMULINK's LINSIM simulates the model of the open loop Flexible Beam [6].

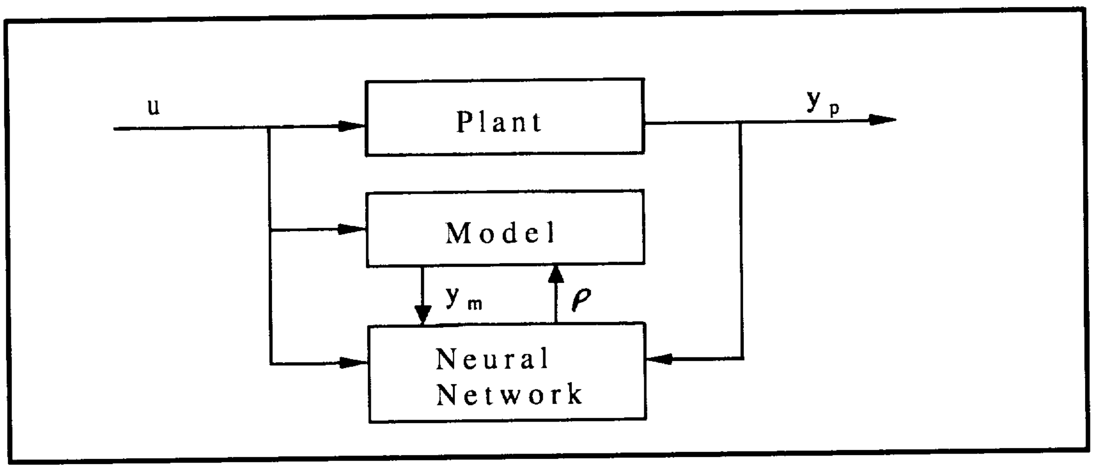


Figure 2. Artificial Neural Network Estimation Training Setup

In applying ANNs one looks for patterns in the data sets. However, sometimes the patterns are not readily apparent, especially over large spatio-temporal ranges. Therefore, limited ranges of spatial and temporal data are investigated and used to perform damping coefficient estimation. Methods of training artificial neural networks are legion (see [9]). However, to perform the training for this problem, a standard gradient descent method with learning rate and momentum adjustments is used. The ANN inputs are the beam sensor outputs and the actuator input for a given simulator run. Training error is the mean squared error of the difference between actual versus simulated flexible beam behavior over the given spatial and/or temporal range.

The input to the SIMULINK model comes from a signal generator, which is assumed to replicate the input of flexible beam actuator. The model output is extracted from the MATLAB Workspace to which it is saved during simulation. The outputs of elements 5, 20, and 40 are stripped from the output and compared with the actual beam sensor outputs to determine the errors which are then used to update neural network weights from which new damping coefficients are computed prior to the next run through the simulation. All inputs to the neural network, including the target data inputs, are normalized using the MATLAB nornmr/c functions [6]. The ANNs consisted of 10 nodes in the first hidden layer, 5 nodes in the second layer, and the number of output nodes equal to the number of element damping coefficients being estimated (in this case 3).

The size of the artificial neural networks is derived from other experiments on training convergence versus network size (see [3]).

Estimation is looked at from three perspectives: spatial, temporal, and spatial and temporal. First, the nets are trained for the same sample time for differing numbers of input frequencies (spatial). The thought is if the nets are trained over many frequencies, the coefficients would be more robust at the given time instant; thereby allowing better adaption at the given timeframe. This is done due to the lack of multiple run data at a single frequency from the actual beam. Sample times are 5, 10, 25, 50,75, 100, 250, and 500. More discretization of the sample space would lead to better performance in a regulator design; however, the objective is proof of concept.

From the temporal perspective the nets are trained over a range of sample times for a single frequency input. This is to see if there is a span of temporal usefulness for the damping coefficients. Spans looked at are 3, 5, 10, 20, 30, and 50 sample times. Finally, both are combined where the training examplars are comprehensive sets of spatial and temporal information.

For the Flexible Beam data the sample period is 62.66 msec with a peak input force of 0.61 Newtons. 'This is also the simulated sample rate and force input. The estimation evaluation criteria is: when using the estimated damping coefficients in the model, does it perform like the actual flexible beam using beam data previously not seen (by the neural networks) during training.

4. PARAMETER ESTIMATION RESULTS

The artificial neural networks are able to provide usable estimates of the flexible beam damping coefficients. When damping coefficient estimates are input to the model and compared against actual beam behavior for operations both previously and not previously seen by the nets, the model successfully emulates the beam behavior. Representative graphs for one spatio-temporal case, Figures 3-5, show undamped beam behavior for two frequencies covering a ten sample times interval in Series 1 and 2; damped behavior in Series 3 and 4, and actual beam behavior (with a DC component) in Series 5 and 6 for comparison.

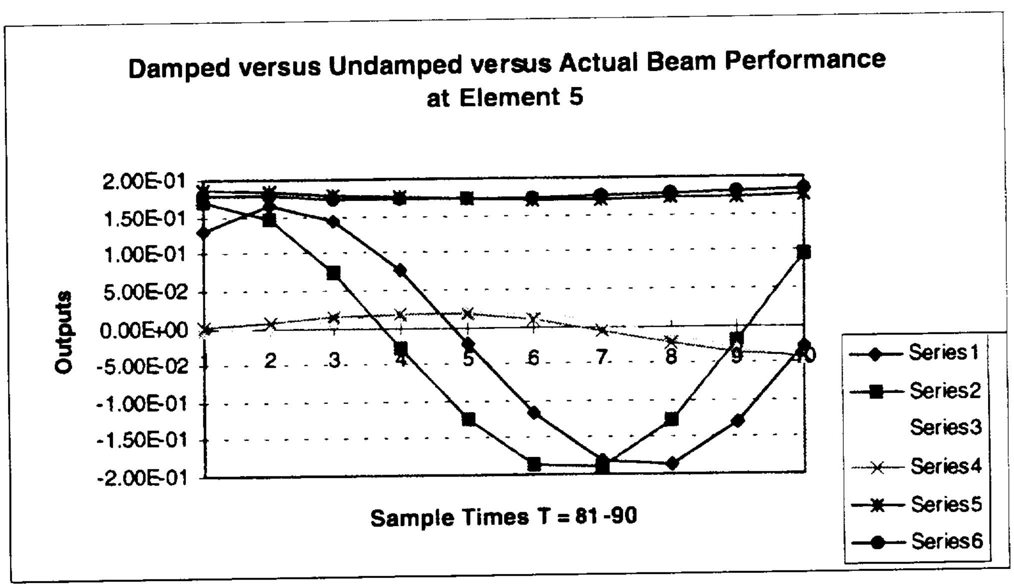


Figure 3. Model versus Actual Beam Behavior for Element 5 at T = 81-90

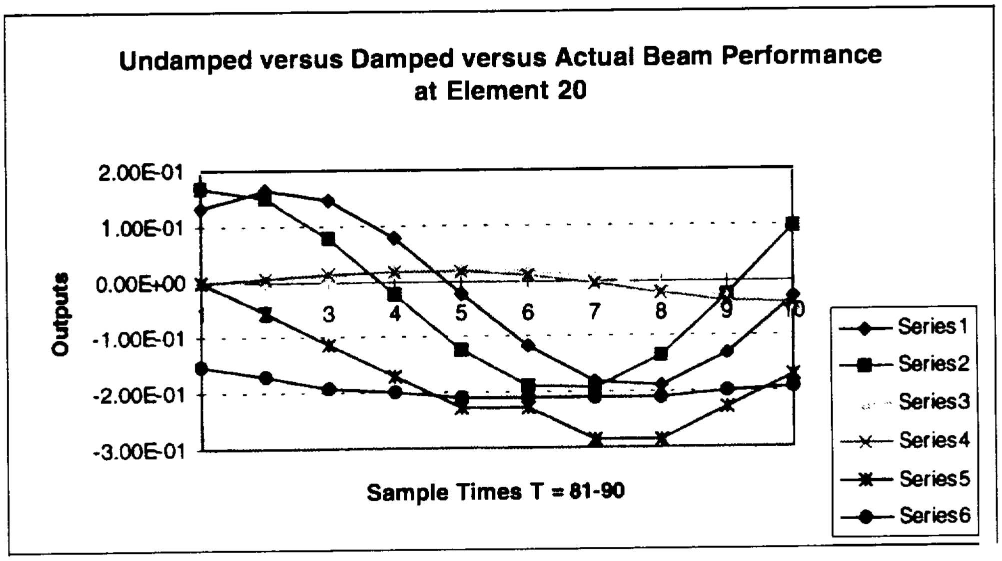


Figure 4. Model versus Actual Beam Behavior for Element 20 at T = 81-90

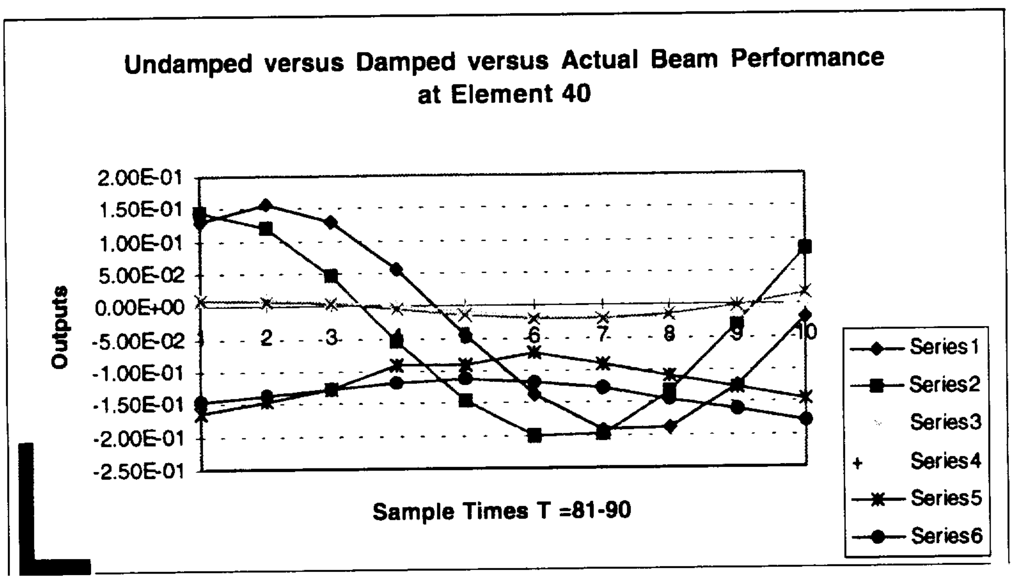


Figure 5. Model versus Actual Beam Behavior for Element 40 at T = 81 -90

Results indicated a sensitivity to the number of temporal sample intervals over which the nets are trained. Either less than a five sample time interval or more than a ten sample time interval yielded poor training and in some cases did not converged to stable estimates. Little sensitivity is observed when the ANNs are trained spatially (multiple frequencies at a single sample time). While reported here, data supporting these observations is found in reference 3. Finally, the components of intelligence portrayed using ANNs in the above application are: adaptability, learning, and dealing with significant complexity [2,10].

CONCLUSIONS AND RECOMMENDATIONS

An intelligent method is shown to provide useful damping coefficient(s) estimates. Artificial neural network training for estimating the damping coefficients is reported to be temporally sensitive and spatially flexible. Results for a representative spatio-temporal instance are provided to show the ability of the intelligent parameter estimator to provide useful damping coefficient estimates over a given time interval of the open loop flexible beam operation.

Future focus will be on characterization of the initiation of motion (T=1-100) as the ability to control the beam should logically be obtained by early initial damping of a given disturbance to the beam. In addition, more actual flexible beam data, especially multiple data sets at individual frequencies is to be obtained and used. Finally, other methods reported in literature or derived by the authors are to be investigated.

REFERENCES

[1] M. Jamshidi, C. Nguyen, R. Lumia, and J. Yuh., Intelligent Automation and Soft Computing: Trends in Research, Development, and Applications, Vol 1, Back Cover, TSI Press, Albuquerque, NM, 1994.

[2] M. A. Johnson, Intelligent Parameter Estimation, EECE 595 Research Project Report, Part I, University of New Mexico, Albuquerque, NM, Mar 96.

[3] M. A. Johnson and E. Hamke, Intelligent Parameter Estimation: Artificial Neural Network and Fuzzy Logic Applications, EECE 595 Research Project Report, Part II and III, University of New Mexico, Albuquerque, NM, Mar 96.

[4] D. P. Peterson, JPL/NASA Flexible Beam Testbed: Analytic and Computational Model and Transducer Interface, Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM, Jun 1994.

[5] P. S. Maybeck, Stochastic Models, Estimation, and Control, Academic Press, Inc., San Diego, CA, 1979.

[6] MATLAB and SIMULINK, The Mathworks, Boston, MA.

[7] FULDEK, Bell Helicopter, Dallas, TX.

[8] M. A. Johnson, and M. B. Leahy, Jr. Adaptive Model-Based Neural Network Control. Proc, 1990 Intl Conf on Robotics and Automation, Vol 3, pp 1704-1711, IEEE Computer Society Press, Los Alamitos, CA, 1990.

[9] M. A. Johnson. Payload Invariant Control via Neural Networks: Development and Experimental Evaluation. Master's Thesis, Air Force Institute of Technology, Air University, October 1989.

[10] M. A. Johnson. Intelligent Methods of Parameter Estimation for Adaptive Control, Proc, Second World Automation Conference, Montpelier, France, May 1996.

[11] K. Kobayashi, Ka. C. Cheok, and K. Watanabe. Fuzzy Logic Rule-Based Kalman Filter for Estimating True Speed of a Ground Vehicle. Automation and Soft Computing, Vol 1, No 2, pp 179-190, Autosoft Press, Albuquerque, 1995.

[12] W. G. McLean and E. W. Nelson. Schaums Outline Series, Theory and Problems of Engineering Mechanics. 4th ed. McGraw Hill Book company, NY, 1988.

URC97066

INTELLIGENT PARAMETER ESTIMATION OF NASA/JPL FLEXIBLE BEAM DAMPING COEFFICIENTS: FUZZY LOGIC KALMAN FILTER

MARK A. JOHNSON
ERIC. HAMKE

NASA Center for Autonomous Control Engineering (ACE)
Department of Electrical and Computer Engineering
University of New Mexico, Albuquerque, NM 87131
Email: ace@pajarito.unm. edu

ABSTRACT

NASA/JPL Flexible Beam damping coefficients are estimated by Fuzzy Logic Kalman Filter (FLKF). Flexible beam damping coefficients are estimated from the difference between actual and modeled open loop beam behavior. Spatial, temporal, and spatio-temporal coefficient estimation is performed. Dynamically estimated damping parameters enable the simulated flexible beam behavior to emulate actual beam behavior where analytically derived static coefficients fail. Results indicate the concept applicability and provide a basis for future work.

1. INTRODUCTION, MOTIVATION, AND APPROACH

Control of structures in space is a critical problem to resolve as it is applicable to anything planned for space stations (orbital or surface), satellites, and inter/intro planetary/stellar vessels. One key to controlling these structures is knowledge of system parameters as they undergo spatio-temporal changes. Parameter estimation is used to provide useful information on important, indirectly observable system parameters. One goal in parameter estimation is to reduce computational complexity while achieving acceptable performance. Intelligent methods provide estimates which are at least as good as conventional analytical methods while reducing computational complexity and estimation delay. Current core intelligent methods are artificial neural networks, fuzzy logic, genetic algorithms, genetic programming, expert systems, and probabilistic reasoning [1]. Performing parameter estimation with potential direct application to adaptive flexible space system(s) control using intelligent techniques is the focus of this paper.

Intelligent methods of estimating parameters are a key component to controlling any future flexible space system. However. little exists in literature concerning experimental applications of intelligent approaches to parameter estimation, which motivates experimentally applying an intelligent method to estimate parameters. A brief synopsis of a portion of the work contained in references 2 and 3 is presented in the following sections. Section 2 begins with the problem statement, followed by the approach taken to solve the problem using fuzzy logic being presented in section 3. Section 4 gives and discusses some experimental results after which conclusions, recommendations, and future directions are presented.

2. PROBLEM STATEMENT

The experimental test bed is graphically depicted in Figure 1. The problem statement is, as follows:

Given: 1) A "distributed-parameter, infinite-degree-of-freedom "plant''"
(NASA/ACE laboratory, NASA/JPL Flexible Beam testbed) [4];
2) A fifty element model of the flexible beam;
3) Two "levitators" at elements 20 and 40 which counterbalance gravity effects and are coupled to optical encoders; and
4) Data from actuator inputs (at element 6) and transducer outputs (at elements 5,20, and 40) during actuator excitation;

Estimate, using intelligent methods, the damping coefficients from actuator and sensor inputs.

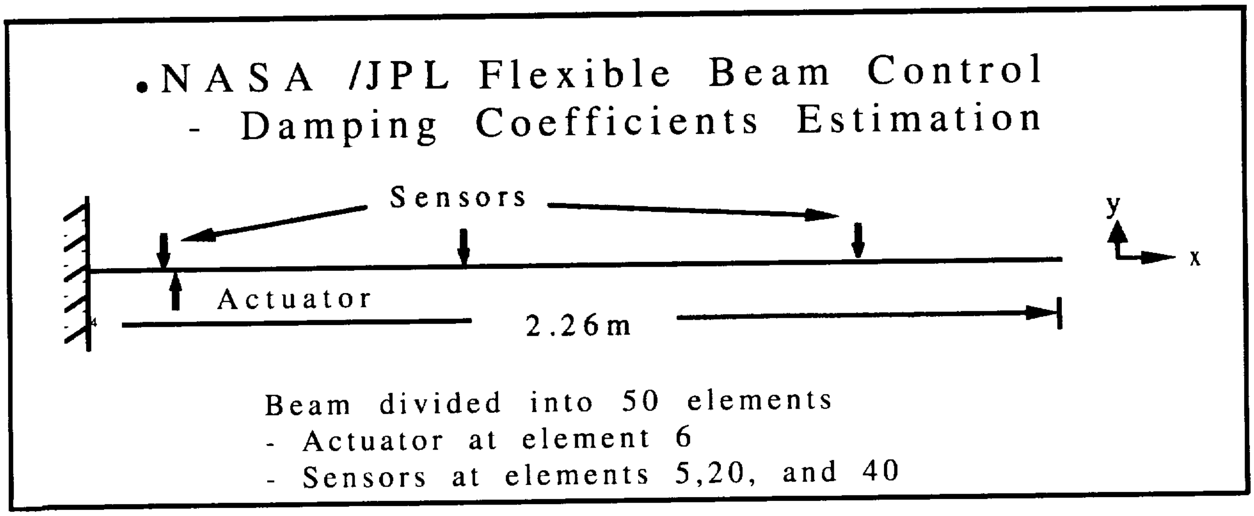


Figure 1. NASA/ACE Flexible Beam Testbed

3. INTELLIGENT PARAMETER ESTIMATION

The basic parameter estimation problem consists of five (5) components [5]. The five components are: 1) The variable(s) to be estimated; 2) Available measurements or observations; 3) Mathematical model describing how the measurements are related to the variable(s) being estimated; 4) Mathematical model of the uncertainties present; and 5) Performance evaluation criteria. Variable(s) to be estimated are the damping coefficients for elements 5, 20, and 40. Each is used due to having sensors on the beam at these location from which direct measurements are made. Only elements 20 and 40 are under direct suspension in an attempt to counteract the effects of gravity. Element 5 is next to element 6, the actuator. One of the advantages of using intelligent techniques is the nonparametric nature of the solution. In this sense the mathematical models of components 3 and 4 are not necessary (see [4] for mathematical model derivation). However, when using Fuzzy Logic, one usually needs some information about how the system will respond to given inputs in order to develop fuzzy rules. Previous experience gained using artificial neural networks [ref?] yields the information needed to appropriately develop rudimentary fuzzy rules.

To understand the evaluation criteria for the problem one frost needs to understand the solution approach. Damping coefficients and their variations within the Flexible Beam structure are not observable. Flexible Beam and model inputs/outputs are observable and available. The hypothesis is to make the model behave like the actual open loop Flexible Beam by dynamically adjusting the model damping coefficients. As the error between the model versus Beam behavior goes to zero, the damping coefficients should converge to their appropriate values since they are the only parameters changing within the model.

Velocity and acceleration measurement estimates taken from standard Kalman Filters are used as inputs to a Fuzzy Logic Kalman Filter (FLKF) implementation by Kobayashi, et al; to estimate the true velocity of a vehicle under skid and slip conditions [11]. Fuzzy rules are used to update measurement covariances based upon understanding which sensor measurements to trust in given situations. In the estimation application being considered the FLKF predicts the damping coefficients from the previous sample time sensor measurements (actual and modeled), and actuator input.

The Kalman Filter is designed in MATLAB using the Control Systems Toolbox so that it may be directly compared with the MATLAB beam model and ANN estimator [3]. Figure 2 depicts the setup for the Fuzzy Logic Kalman Filter (FLKF) parameter estimator. The goal of damping coefficient estimation is to be able use the estimates in a controller which will control disturbances to the beam. Therefore, using reason and neural network experimentation results, the following list is a rudimentary set of fuzzy rules used in the FLKF experiments. The fuzzy rules for updating the R matrix based upon output errors, areas follows:

If $|y_m - y_p|$ is Very Small (VS), Then R is S; or
If $|y_m - y_p|$ is Small (S), Then R is S; or
If $|y_m - y_p|$ is Medium Small (MS), Then R is M; or
If $|y_m - y_p|$ is Medium (M), Then R is M; or
If $|y_m - y_p|$ is Medium Large (ML), Then R is M; or
If $|y_m - y_p|$ is Large (L), Then R is L; or
If $|y_m - y_p|$ is Very Large (VL), Then R is L,

where, S = 0.0002, M = 0.0005, L = 0.0008, y_m is the model output, and y_p is the actual beam output.

These rules, applied individually to each sensor output, cover a multitude of situations. For instance, at the start of a disturbance occurring at element 6 (the actuator), the sensor at element 5 being the closest would yield the most information concerning the disturbance and due to the temporal lag from one end of the beam to the other would initially have a large model versus actual beam output difference. Therefore, the value of R for element 5 would be chosen to be large (more important) than the values for elements 20 and 40.

The following initial values are used to develop the various L, P, Q, and R matricies for the FLKF and are assumed to have zero mean Gaussian distributions. The normal covariance of system noise is set to: Q = 0.0002; and normal covariance of the measured output (sensor) noise is set to: R_s = 0.0005; and the normal covariance of the actuator input noise is set to: R_u = 0.0002. The initial system state covariance is set to $P = BQB^T$. Using the A, B, C, and D matricies the Kalman gain matrix is derived using the MATLAB function 'lqe'. The L matrix returned from 'lqe' is then applied to the Kalman Filter (using MATLAB 'estim' function) for a given input signal

to estimate the damping coefficients for input to the model. The error of the model outputs versus the actual beam outputs for a given timeframe is then used to change the R matrix based upon the appropriate fuzzy rules applied individually to each sensor output.

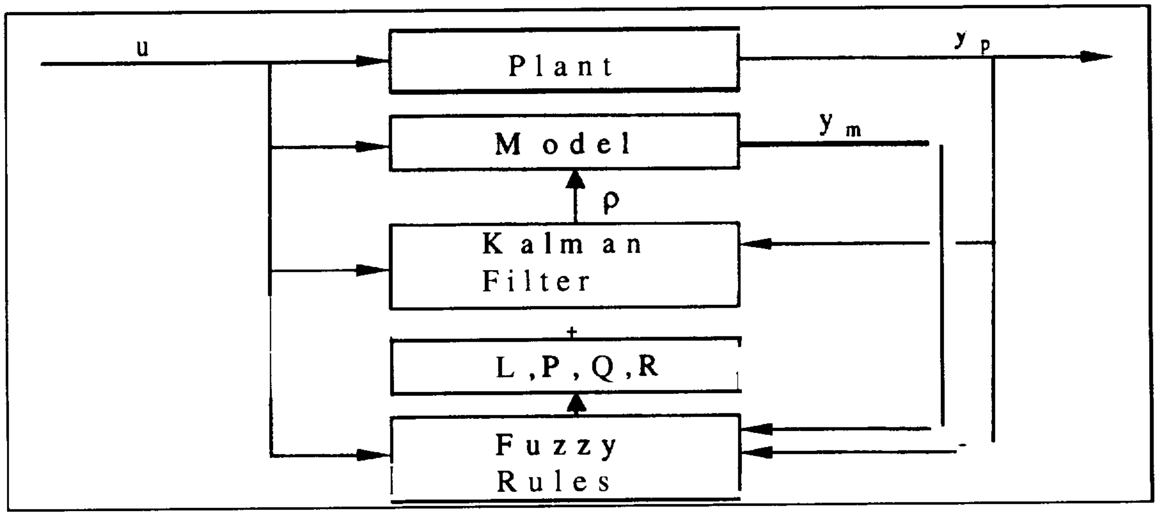


Figure 3. Fuzzy Logic Kalman Filter Experimental Setup

For the Flexible Beam data the sample period is 62.66 msec with a peak input force of 0.61 Newtons, which is also the simulated sample rate and force input. The estimation evaluation criteria is: when using the estimated damping coefficients in the model, does it perform like the actual flexible beam.

4. FLKF PARAMETER ESTIMATION RESULTS

The Fuzzy Logic Kalman Filter to provides usable estimates of the flexible beam damping coefficients. When damping coefficient estimates are input to the model and compared againstactual beam behavior the model reasonably emulates the beam behavior. Figures 4-6, show undamped model beam behavior for two frequencies covering a thirty sample times interval (T=235-265) in Series 1 and 2; damped model behavior in Series 3 and 4, and actual beam behavior in Series 5 and 6 for comparison. The damping coefficients for elements 5, 20, and 40 are: -28.4123,-7.7297, and 1.2356, respectively. The mean squared error is about 0.04. Towards the end of the sample interval, model and actual beam behavior appear to diverge. Further work is needed to determine whether or not it is due to the limits placed on the possible gain values (i.e. 0.0002 and 0.0008).

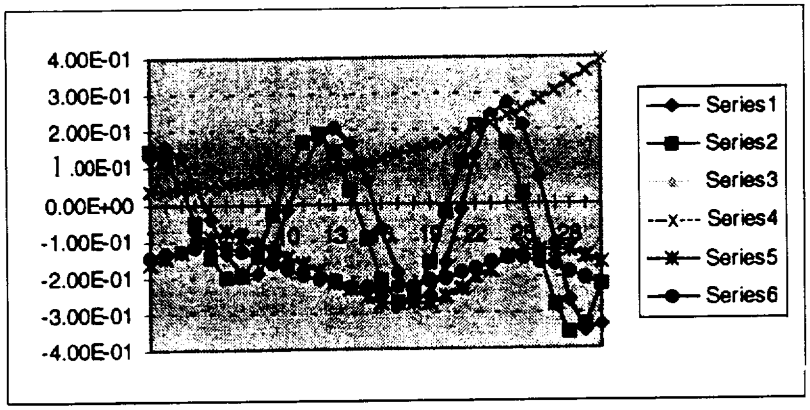


Figure 3. Model versus Actual Beam Behavior for Element 5 at T = 235-265

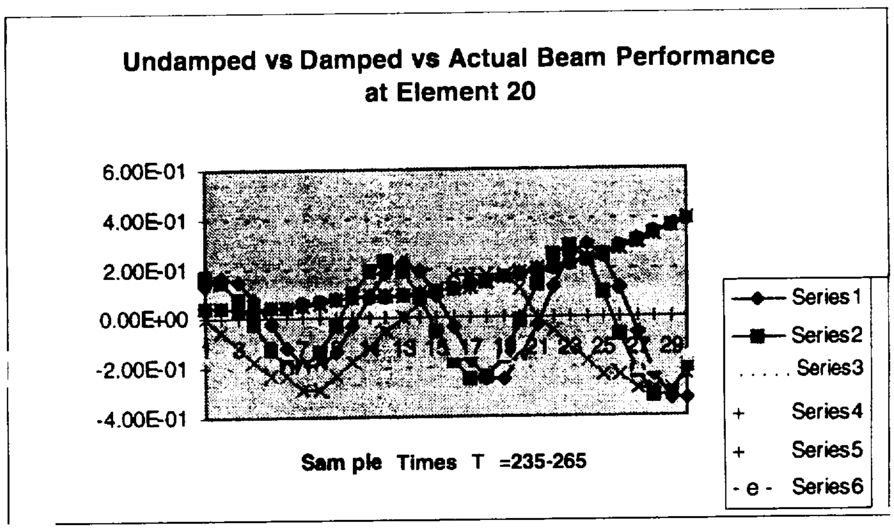


Figure 4. Model versus Actual Beam Behavior for Element 20 at T = 235-265

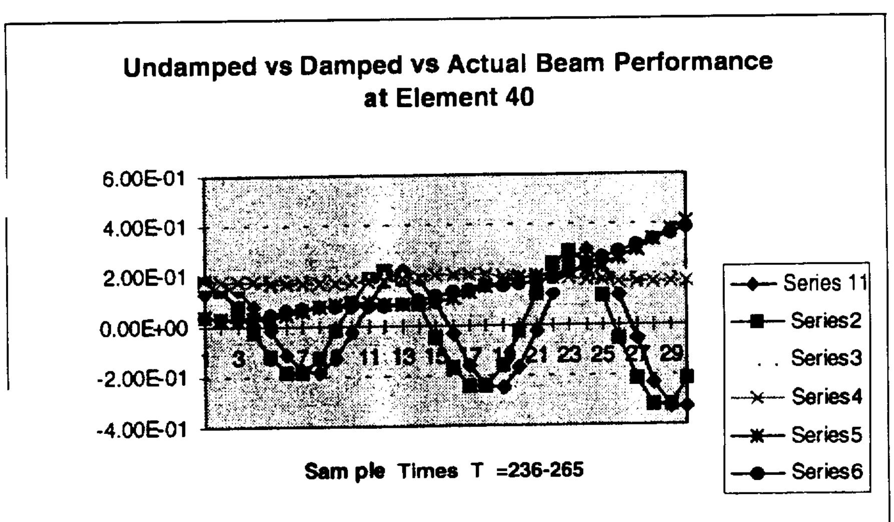


Figure 5. Model versus Actual Beam Behavior for Element 40 at T = 235-265

CONCLUSIONS AND RECOMMENDATIONS

A Fuzzy Logic Kalman Filter is shown to provide useful damping coefficient(s) estimates. Results for a representative spatio-temporal instance are provided to show the ability of the intelligent parameter estimator to provide useful damping coefficient estimates over a given time interval of the open loop flexible beam operation.

Future focus will be on characterization of the initiation of motion (T= 1- 100) as the ability to control the beam should logically be obtained by early initial damping of a given disturbance to the beam. In addition, more actual flexible beam data, especially multiple data sets at individual frequencies is to be obtained and used. Finally, other methods reported in literature or derived by the authors are to be investigated.

REFERENCES

[1] M. Jamshidi, C. Nguyen, R. Lumia, and J. Yuh., Intelligent Automation and Soft Computing: Trends in Research, Development, and Applications, Vol 1, Back Cover, TSI Press, Albuquerque, NM, 1994.

[2] M. A. Johnson, Intelligent Parameter Estimation, EECE 595 Research Project Report, Part I, University of New Mexico, Albuquerque, NM, Mar 96.

[3] M. A. Johnson and E. Hamke, Intelligent Parameter Estimation: Artificial Neural Network and Fuzzy Logic Applications, EECE 595 Research Project Report, Part II and III, University of New Mexico, Albuquerque, NM, Mar 96.

[4] D. P. Peterson, JPL/NASA Flexible Beam Testbed: Analytic and Computational Model and Transducer Interface, Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM, Jun 1994.

[5] P. S. Maybeck, Stochastic Models, Estimation, and Control, Academic Press, Inc., San Diego, CA, 1979.

[6] MATLAB and SIMULINK, The Mathworks, Boston, MA.

[8] M. A. Johnson, and M. B. Leahy, Jr. Adaptive Model-Based Neural Network Control. Proc, 1990 Intl Conf on Robotics and Automation, Vol 3, pp 1704-1711, IEEE Computer Society Press, Los Alamitos, CA, 1990.

[9] M. A. Johnson. Payload Invariant Control via Neural Networks: Development and Experimental Evaluation. Master's Thesis, Air Force Institute of Technology, Air University, October 1989.

[10] M. A. Johnson. Intelligent Methods of Parameter Estimation for Adaptive Control, Proc, Second World Automation Conference, Montpelier, France, May 1996.

[11] K. Kobayashi, Ka. C. Cheek, and K. Watanabe. Fuzzy Logic Rule-Based Kalman Filter for Estimating True Speed of a Ground Vehicle. Automation and Soft Computing, Vol 1, No 2, pp 179-190, Autosoft Press, Albuquerque, 1995.

URC97067

N/A
567/19

Structural Analysis of the QCM Aboard the ER-2•

Phyllis D. Jones, Peter M. Bainum, and Guangqian Xing
Dept. of Mechanical Engineering, Howard University, Washington, DC 20059

Introduction

As a result of recent supersonic transport (SST) studies on the effect they may have on the atmosphere, several experiments have been proposed to capture and evaluate samples of the stratosphere where SST's travel. One means to achieve this is to utilize the quartz crystal microbalance (QCM) installed aboard the ER-2, formerly the U-2 reconnaissance aircraft. The QCM is a cascade impactor designed to perform in-situ, real-time measurements of aerosols and chemical vapors at an altitude of 60,000- 70,000 feet. The ER-2 is primarily used by NASA for Earth resources to test new sensor systems before they are placed aboard satellites. One of the main reasons the ER-2 is used for this flight experiment is its capability to fly approximately twelve miles above sea level (can reach an altitude of 78,000 feet)[1]. Because the ER-2 operates at such a high altitude, it is of special interest to scientists interested in space exploration or supersonic aircraft. Some of the experiments are designed to extract data from the atmosphere around the ER-2. For the current flight experiment, the QCM is housed in a frame that is connected to an outer pod that is attached to the fuselage of the ER-2. Due to the location of the QCM within the housing frame and the location of the pod on the ER-2, the pod and its contents are subject to structural loads. In addition to structural loads, structural vibrations are also of importance because the QCM is a frequency induced instrument. Therefore, a structural analysis of the instrument within the frame is imperative to determine if resonance and/or undesirable deformations occur.

Structural Analysis

In order to perform a structural analysis of the QCM within its housing frame, a finite element mathematical model of the system (QCM and housing frame) was created. Initially, separate finite element models of the QCM and housing frame were constructed, Figures I & II respectively. Once the individual models were produced, they were graphically incorporated to create the system as shown in Figure III. The models were graphically generated by the finite element graphical software package MSC/PATRAN[2]. By inputting the structure's geometry, material properties, inertia, and boundary conditions, MSC/PATRAN has the capabilities to generate a finite element model. The geometry consists of the physical make-up of the structure and its dimensions with respect to a given. origin. For the QCM, the dimensions are length=1 8.3 in., width=4.5 in., and height=4.8 in. and for the housing frame the length=66.5 in., width=1 1.5 in., and height=13.2 in. For this system, both the QCM and housing frame are composed of aluminum with material properties of 10 E6 $\frac{lbs}{in^2}$ for Young's modulus, 0.333 for Poisson's ratio, and 0.98 $\frac{lbs}{in^3}$ for the density. The QCM rests lengthwise on an aluminum plate within the frame and is rigidly bounded at both ends, see Figure III. Once the system is graphically generated, the finite element model data is transferred into MSC/NASTRAN which is utilized to generate analytical results, see Table I, based on the data from MSC/PATRAN. After the analytical results are compiled, the results are input back into MSC/PATRAN for further graphical analysis. As an example, the various mode shapes for the system are depicted in Figure IV.

• Research supported by the NASA/Howard University Center for the Study of Terrestrial & Extraterrestrial Atmospheres (CSTEA)

MODE NUMBER	CYCLIC FREQUENCY
1	130.159
2	208.456
3	268.106
4	282.352
5	324.958
6	338.102
7	355.453
8	359.442
9	360.811
10	388.971

Table I: Analytical Results

Results

The results indicate several of the mode shapes of the system could cause errors in the reading of the QCM, see Figure IV, due to excessive deflections associated with several of the vibrational modes. Within the housing frame, the QCM is supported by an aluminum plate. Certain mode shapes indicate that this plate and the QCM could have large distortions (deflections) which directly effect the readings of the QCM.

The QCM contains quartz crystals which are used to accumulate air sample particulates[3]. Once these samples are collected, the instrument records the change in frequency between the reference crystal and the crystal that collects the air samples. Because the QCM relies on changes in frequency, the deflections could cause interference with these changes in the frequency readings. One means of minimizing or eliminating the deflections would be to attach a vibration absorber (modeled by a) spring-mass-damper to a specific point(s) on the system where large deflections are occurring[4]. By examining which locations on the system yield the largest deflections that directly affect the QCM, specific nodal point(s) within the system's mathematical model can be chosen as possible points to apply the spring-mass-damper. By determining from the modal participation factors, points of maximum deflection of the eigenvectors and mode shapes of the system, the nodal points of greatest deflections can be determined to yield the best possible points of application for the spring-mass-damper. Once these nodal point(s) are determined, the spring-mass-damper is mathematically designed and then implemented into MSC/NASTRAN. The analytical results are computed and again the results are input back into MSC/PATRAN to determine if the point(s) of application of the spring-mass-damper are successful. This process is repeated until an optimum design is obtained to minimize or eliminate the deflections.

Acknowledgments

I would like to acknowledge and thank Mrs. Sandra Irish, the group leader of the Structural Analysis Section at NASA Goddard Space Flight Center, Greenbelt, MD, for all of her help and time in assistance with this project.

References

[1] Stewart, Doug, "Above the Sky", *Air & Space,* August/September 1993, pp. 11-13.

[2] Miller, Mark, *User's Guide: Getting Started With MSC/NASTRAN, 1st Ed.*, MacNeal - Schwendler Corp., Los Angeles, CA, December 1993.

[3] Jones, Phyllis D., "Description of the QCM/SAW Utilized for the Concorde/ER-2 Flight Experiment", Technical Note, Dept. of Aerospace Engineering Sciences, University of Colorado, Boulder, CO, August 1994.

[4] Marsis, Wisjnu & Bainum, Peter, "Hybrid Control System for Space Mast Structures", AAS/AIAA Astrodynamics Specialist Conference, Durango, CO, August 19-22, 1991 Paper No. AAS -91-374.

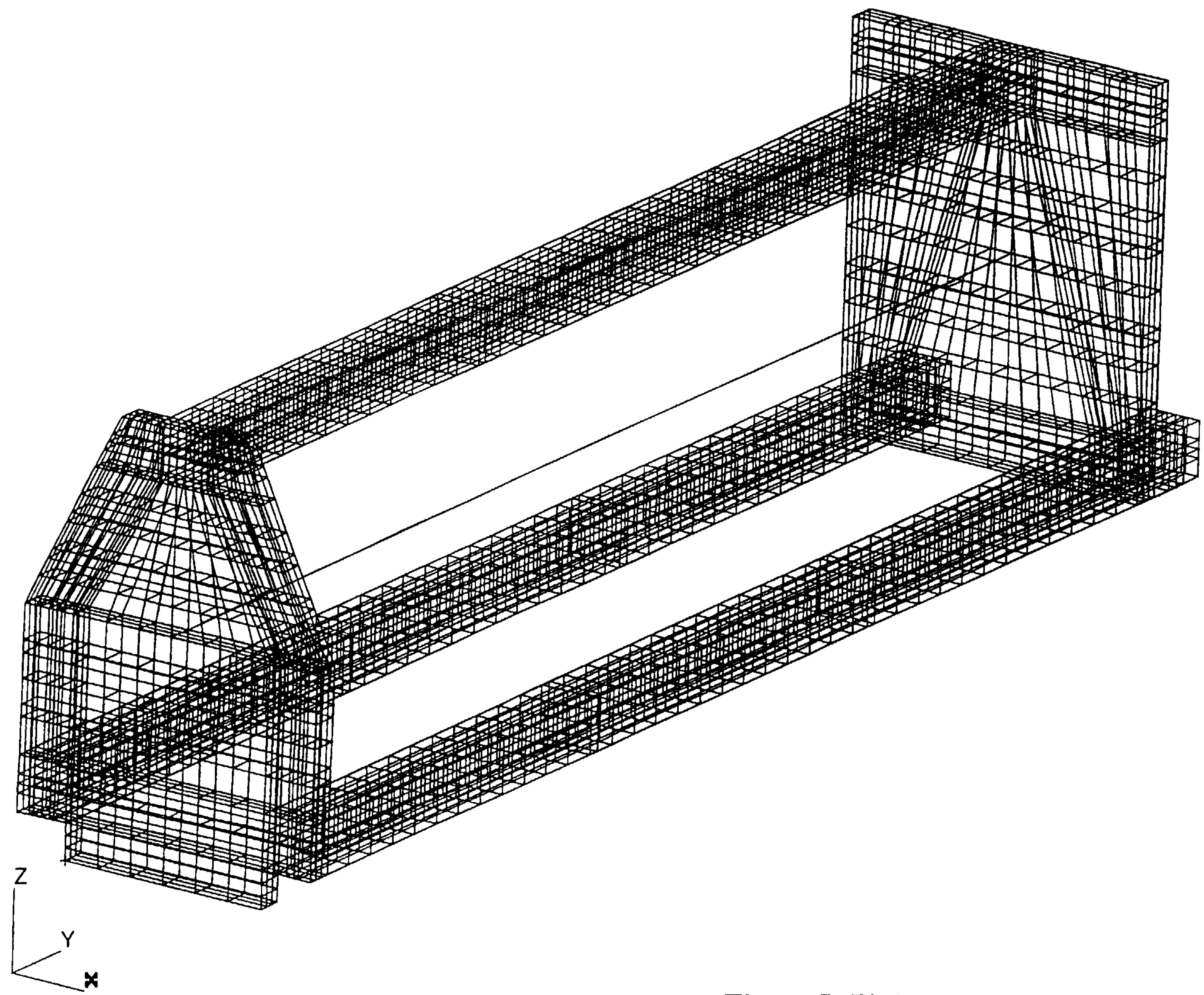


Figure I: Finite Element Model of the QCM

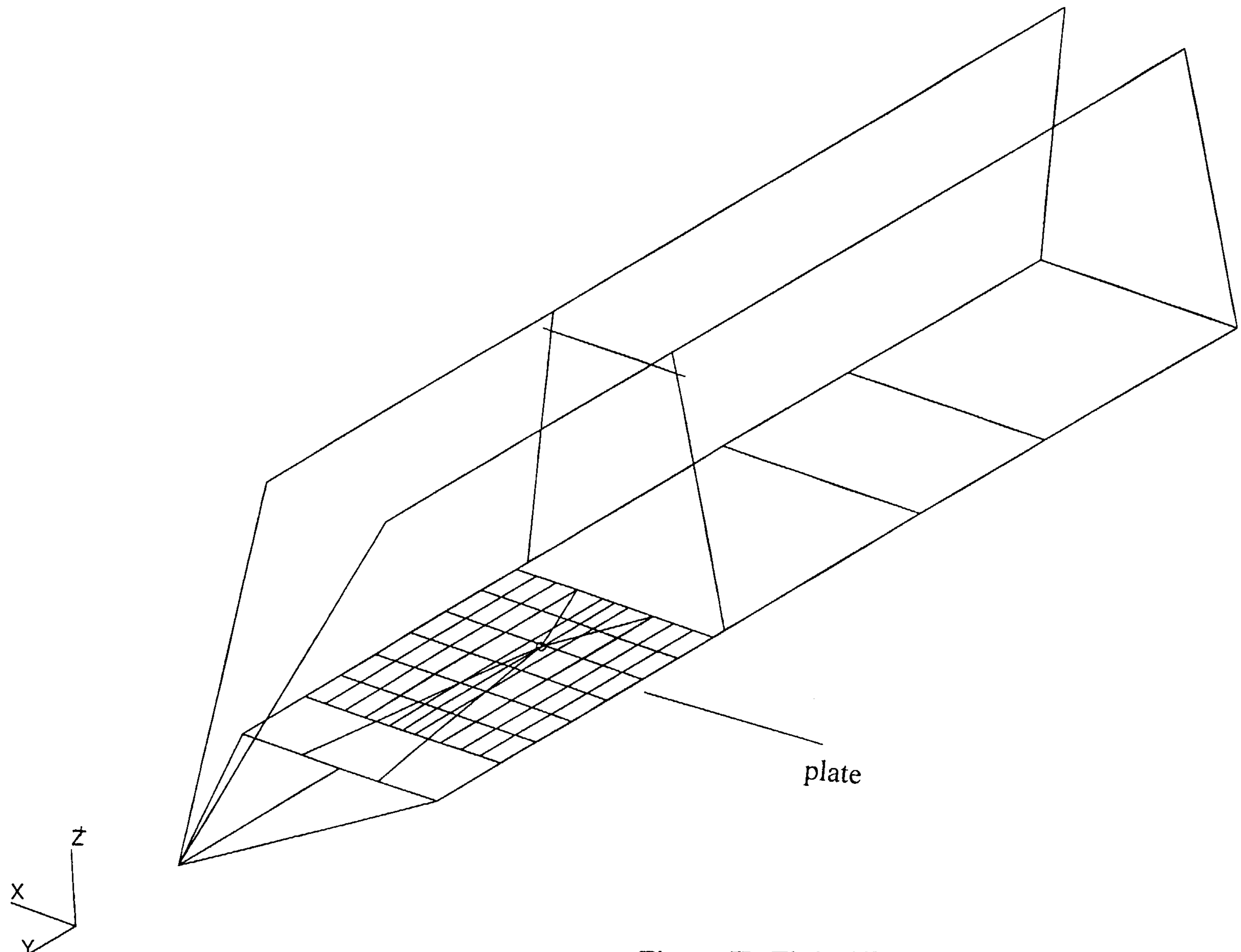


Figure II: Finite Element Model of the Housing Frame

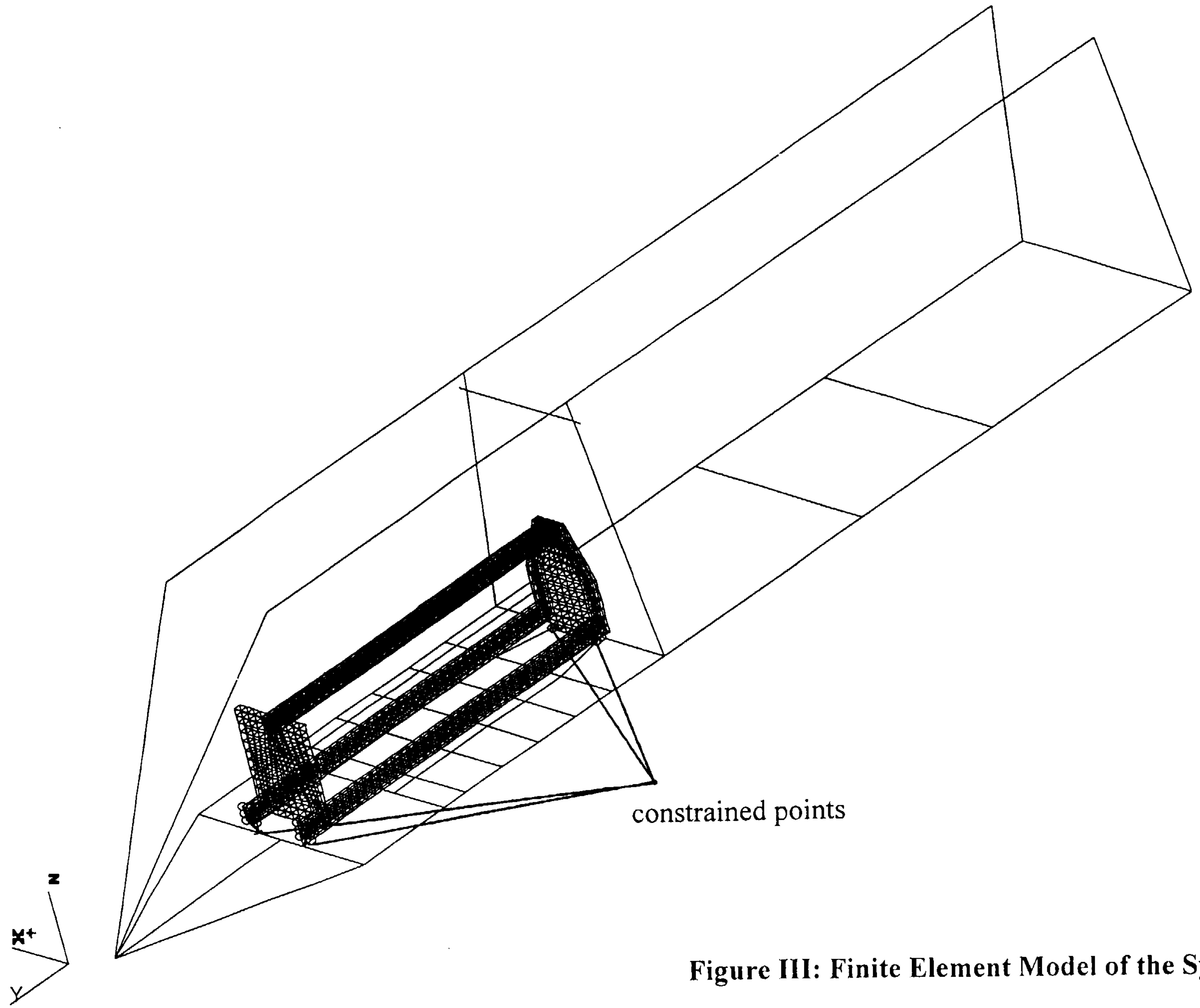


Figure III: Finite Element Model of the System

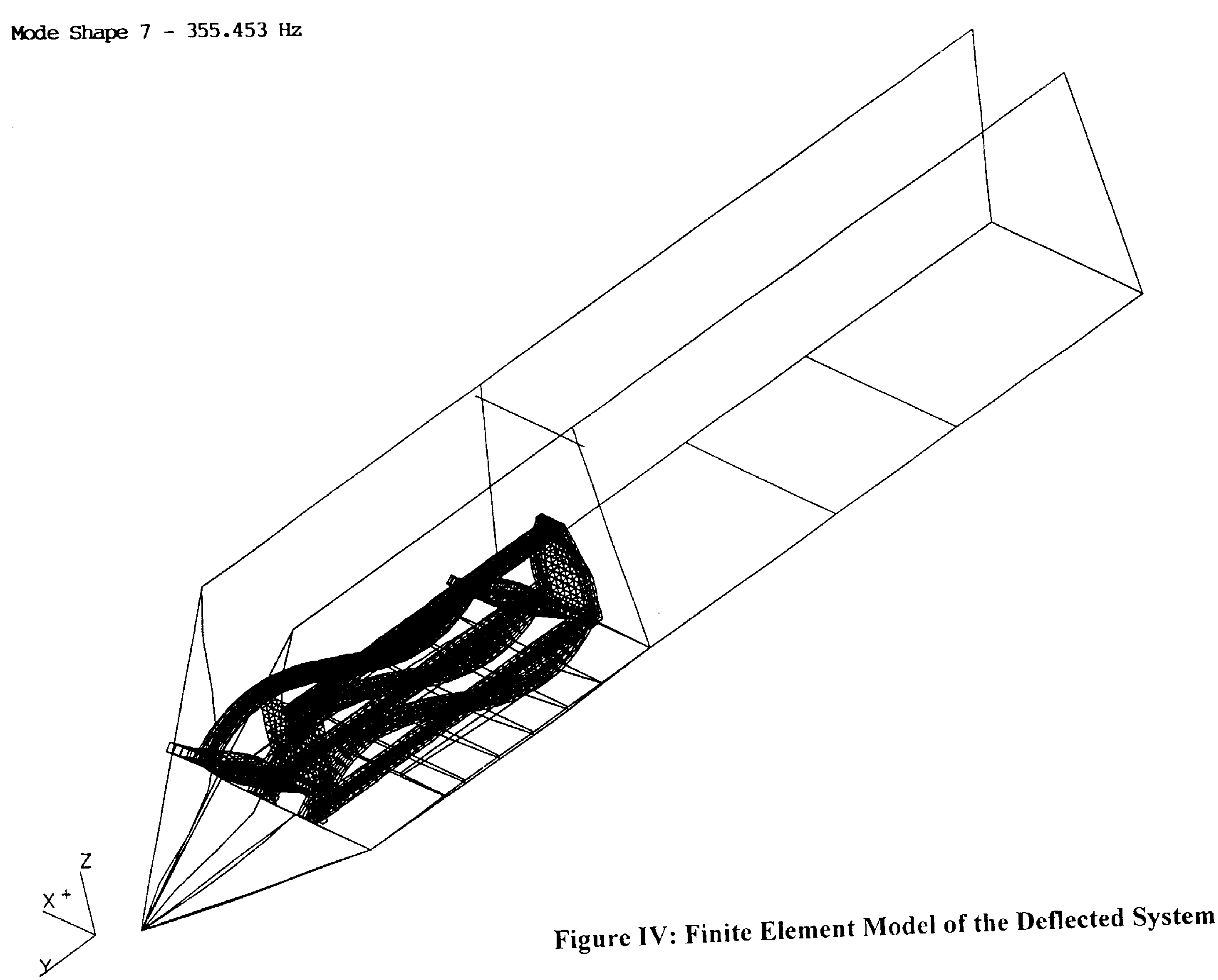


Figure IV: Finite Element Model of the Deflected System

568/25

URC97068

FOURIER TRANSFORM INFRARED (FT-IR) SPECTROSCOPY OF NITROGEN DIOXIDE, SULFUR DIOXIDE, HYDROGEN CHLORIDE, AND METHYL NITRITE PERTAINING TO ATMOSPHERIC PHENOMENA

John Jordan
Department of Electrical Engineering&
Center for the Study of Terrestrial and Extraterrestrial Atmospheres (CSTEA)
Howard University
Washington, DC 20059

Hassan Lauziere, Mohammed Kamal, Chandran Haridass & Prabhakar Misra, Department of Physics and Astronomy&
Hideo Okabe, Department of Chemistry&
Center for the Study of Terrestrial and Extraterrestrial Atmospheres (CSTEA)
Howard University
Washington, DC 20059

ABSTRACT

Fourier Transform Infrared (FT-IR) spectroscopy is a dynamical instrumentation technique that is useful for examination of various forms of matter. When a sample is analyzed using infrared radiation, the FT-IR spectra provide information about the vibrational and rotational energies of a molecule. The infrared radiation is absorbed at specific frequencies by a particular molecular species, whereby the FT-IR spectrum is a signature of the molecule. Knowing the infrared frequency in relation to the radiation intensity, one can specify the types of chemical functional groups present in a particular sample. The analysis of the rotation-vibration spectra provides information about the molecular constants and the temperature of the in frared source.

INTRODUCTION

Nitrogen dioxide (NO_2) (Abina, et al., 1996), sulfur dioxide (SO_2), hydrogen chloride (HCl), and methyl nitrite (CH_3ONO) (Kamal, 1994) are known pollutants in the atmosphere. Nitrogen dioxide is found in emissions from aircraft and is a participant in significant atmospheric photochemistry. Sulfur dioxide is a known toxin in the atmosphere that is emitted from volcanoes and power plants. Hydrogen chloride is a pollutant that also participates in photochemical air pollution. Methyl nitrite is a precursor of the methoxy radical, which is an important chemical intermediate in combustion reactions and in photochemical atmospheric processes. A clearer understanding of these species and the roles these chemicals play in the combustion chain also lead to the development of improved and efficient organic fuels, besides a greater understanding of atmospheric phenomena.

Fourier Transform Infrared (FT-IR) spectroscopy is a particularly useful analytical technique because of its versatility (Coleman, 1993). Spectra can be obtained of samples in all three states of matter and in most cases nondestructively. Its distinct advantage over other forms of spectral data acquisition is that it has the ability to look at all the wavelengths of a spectral region simultaneously rather than one wavelength at a time. Such an approach saves time dramatically and utilizes light more efficiently. FT-IR spectroscopy has been used to record the spectra of atmospherically significant gas phase nitrogen dioxide (NO_2), sulfur dioxide (SO_2), hydrogen chloride (HCl), and methyl nitrite (CH_3ONO).

EXPERIMENTAL

Sample Preparation of the Experimental Gases NO_2, SO_2, HCl, & CH_3ONO

NO_2: 15.2 psi of nitrogen dioxide (NO_2) gradually flowed in an evacuated gas cylinder.

It was later filled with 960.5 psi of dry nitrogen (N_2) to obtain a sample of 1% nitrogen dioxide.

SO_2: ***484.0*** ppm of sulfur dioxide (SO_2) was commercially obtained from Scott Specialty Gases, Inc. This gas was premixed with air.

HCl: **529.0** ppm of hydrogen chloride (HCl) was commercially obtained from Scott Specialty Gases, Inc. This gas was premixed with dry nitrogen (N_2).

CH_3ONO: Methyl nitrite is not available commercially and has to be synthesized. Two solutions were prepared. One solution contained sodium nitrite ($NaNO_2$), distilled water (H_2O), and methyl alcohol (CH_3OH), while the second solution comprised of concentrated sulfuric acid (H_2SO_4) and distilled water (H_2O).

Two low-temperature baths were prepared using a mixture of methanol and liquid nitrogen. The first solution was prepared by mixing the ingredients together in a three-legged flask. 255 mL distilled water was taken in a small beaker. The beaker was kept in the second cold temperature bath and the concentrated sulfuric acid was added very slowly to the water in the beaker, while continuously stirring the cold dilute H_2SO_4. The dilute H_2SO_4 was taken in a cylindrical separation funnel that was connected to the middle leg of the three-legged flask. The first solution was maintained cold by immersion in another dish containing ice and water. Water vapor and CO_2 inside the flask were purged by connecting the right leg of the flask to gaseous nitrogen (N_2). The left leg of the flask was connected to the exhaust system through a glass trap. The stop-cock of the funnel was opened and the cool dilute H_2SO_4 was added to the solution in the flask drop-by-drop.

After the reaction was complete, the entire solution was transferred to a graduated cylinder. Methyl nitrite was transferred to a cylinder immediately after the synthesis, because its boiling point (-18 .0°C) was above room temperature. After preparation, 2 psi of methyl nitrite was mixed in a gas cylinder containing 198 psi of helium to obtain 1% methyl nitrite (CH_3ONO).

Figure 1
Nicolet Magna-IR 550 Fourier Transform Infrared spectrometer with a 10 m Gas Cell

Description of the FT-IR Spectrometer equipped with a 10 meter cell

Infrared spectra were recorded using a Nicolet Magna-IR 550 Fourier Transform Infrared spectrometer as shown in Figure 1. A multipass 10 m absorption cell was used for recording the spectra and showed much improved sensitivity as compared to the traditional 10 cm cell. Two flow methods were used for spectral recording of the samples: quasi-static and free-flowing. Various materials for one meter gas flow tubing were used to verify analysis and examine reactivity of the sample with the tubing medium. Resolution was set at 4 cm^{-1} for the FT-IR spectra recording of all samples. Thirty-two scans were taken every time a sample was examined and the results were superimposed and averaged for final spectral illustration. Sample cells used KBr windows for mid-infrared transmission. All samples were analyzed in gas phase and over a range of pressures. Spectra were observed in the mid infrared range of 4000 to 400 cm^{-1}. Measurements were localized for nitrogen dioxide in the range 1850 to 1230 cm^{-1}. For sulfur dioxide the range was 1425 to 400 cm-1, whereas for hydrogen chloride a range of 3100 to 2600 cm^{-1} was used. The above cited wavenumber ranges provided signature FTIR plots for the different molecular species.

FT-IR Spectrum of NO_2 and HNO_3 using a 10 m Gas Cell

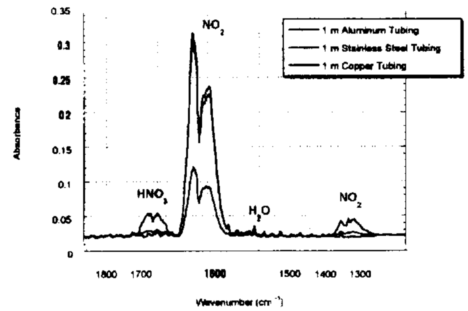


Figure 2
FT-IR spectrum of 1 % nitrogen dioxide (NO_2) in N_2 under quasi-static conditions at 4.0 cm-' resolution with a gas pressure of 200 torr

FT-IR Spectrum of NO_2 and HNO_3 using a 10 m Gas Cell

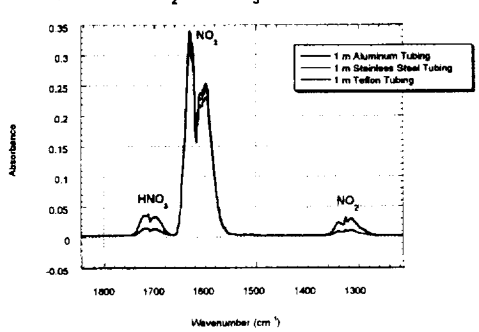


Figure 3
FT-IR spectrum of 1 % nitrogen dioxide (NO_2) in N_2 under free-flowing conditions at 4.0 cm-' resolution with a gas pressure of 200 torr

RESULTS AND DISCUSSION

FT-IR Spectral Analysis of Nitrogen Dioxide

Figure 2 shows the FT-IR spectrum in the spectral range 1850-1230 cm^{-1} of 1% nitrogen dioxide (NO_2) and nitric acid (HNO_3) under dry nitrogen (N_2) recorded at a resolution of 4 cm^{-1} and with a gas pressure of 200 torr under quasi-static conditions using aluminum, stainless steel and copper tubing. Figure 3 shows the FT-IR spectrum of 10/0 nitrogen dioxide (NO_2) in dry nitrogen (N_2) recorded at a resolution of 4 cm-' and with a gas pressure of 200 tom in free-flowing conditions using aluminum, stainless steel and teflon tubing.

Table 1
FT-IR spectral peaks of 1 % nitrogen dioxide (NO_2) in N_2 under quasi-static conditions at a gas pressure of 200 torr using a 10 m gas cell

	Aluminum	Stainless Steel	Copper
HNO_3 1719.0	--------	0.0538	0.0296
HNO_3 1712.5	--------	0.0540	0.0296
HNO_3 1699.6	--------	0.0537	0.0305
NO_2 1628.8	0.301	0.315	0.120
NO_2 1608.3	0.226	0.224	0.0898
NO_2 1596.4	0.224	0.237	0.0916
NO_2 1339.3	--------	0.0422	--------
NO_2 1324.9	--------	0.0457	--------
NO_2 1313.9	--------	0.0451	--------

Table 2
FT-IR spectral peaks of 1 % nitrogen dioxide (NO_2) in N_2 under free-flowing conditions at a gas pressure of 200 torr using a 10 m

	Aluminum	Stainless Steel	Teflon
HNO_3 1719.0	--------	0.0365	0.0296
HNO_3 1712.5	--------	0.0385	0.0315
HNO_3 1699.6	--------	0.0336	0.0263
NO_2 1628.8	0.311	0.328	0.275
NO_2 1608.3	0.217	0.229	0.192
NO_2 1596.4	0.231	0.244	0.203
NO_2 1339.3	--------	0,0241	0.0192
NO_2 1324.9	--------	0.0321	0.0263
NO_2 1313.9	--------	0.0301	0.0248

Notice the different measurement responses with the various tubing in Figures 2 & 3. Stainless steel shows the best absorbance response in comparison to the other tubing.

Likewise, teflon tubing shows a very good absorbance response under free-flowing condition. Aluminum is not as sharp as the other two materials (in both figures). Copper tubing shows a poor absorbance response in comparison to all the tubing observed. Nitrogen dioxide is a very corrosive gas and tends to have high reactivity with other metals. This is evident in the case of copper and aluminum tubing. Stainless steel is a corrosive protestant, while teflon is a plastic-like material that is fairly inert. The corrosive nature of nitrogen dioxide thus shows lower reactivity with these two materials, namely stainless steel and teflon.

A weak feature around 1500 cm^{-1} in Figure 2 due to O-H bending (ν_2).conforms the presence of water vapor. It is present in the spectra recorded under quasi-static conditions for nitrogen dioxide. These weak feature is not observed under free-flowing conditions as shown in Figure 3. The free-flowing gas flow method thus eliminates unwanted anomalies like water (H_2O) and allows for better spectral analysis.

The absorption intensities of various molecular species along with the wavenumbers for 1% nitrogen dioxide (NO_2) under quasi-static and free-flowing conditions are presented in Tables 1 &2.

In Table 1 one can notice that NO_2 reacts with copper more than the other two materials. This is confined by the low values of absorption intensities in the ν_1 mode of NO_2 around 1600 cm-l.

FT-IR Spectral Analysis of Sulfur Dioxide

Figure 4 shows the FT-IR spectrum in the spectral range 1425 to 400 cm^{-1} of sulfur dioxide (SO_2) in air recorded at a resolution of 4 cm-' and with a gas pressure of 400 torr under quasi-static conditions using aluminum, stainless steel, copper, and teflon tubing. Table 3 summarizes the FT-IR spectral features in the range 1425 to 400 cm^{-1} of sulfur dioxide (SO_2) in air recorded at a resolution of 4 cm^{-1} and with a gas pressure of 400 tom under quasi-static condition.

Notice in Figure 4 that the spectra for all the tubing overlap one another. This shows that reactivity of sulfur dioxide with aluminum, stainless steel, copper, and teflon are the same. Table 3 shows the spectral peaks of sulfur dioxide. The spectral peaks for sulfur dioxide are exactly the same in all observed material tubing. The spectral peaks observed in Table 3 represent the spectral peaks of all tubing observed in Figure 4.

FT-IR Spectrum of SO_2 using 10 m Gas Cell

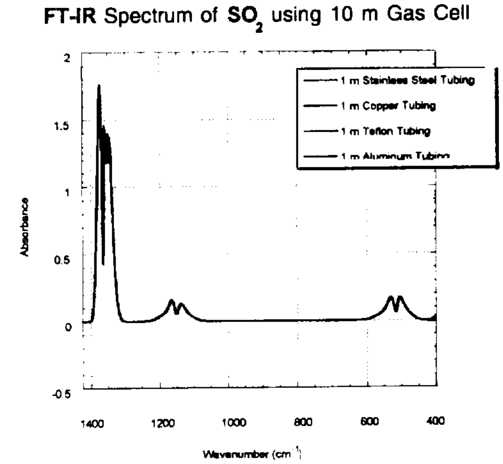


Figure 4
FT-IR spectrum of sulfur dioxide (SO_2) in air under quasi-static conditions at 4.0 cm-l resolution with a gas pressure of 400 torr

Table 3
FT-IR spectral peaks of sulfur dioxide (SO_2) in air under quasi-static conditions at a gas pressure of 400 torr using a 10 m Gas Cell

ν (cm^{-1}) (cal)	ν(cm^{-1}) (obs)	Mode	Absorbance
1361.0	1372.58	$b_1(SO_2)$	1.762
	1358.80		1.529
	1349.42		1.408
	1345.02		1.392
1151.2	1165.10		0.160
	1136.75	$a_1(SO_2)$	0.133
519	530.75	$a_1(SO_2)$	0.172
	504.99	"	0.171

FT-IR Spectral Analysis of Hydrogen Chloride

Figure 5 shows the FT-IR spectrum of the fundamental band (1 -O) of the diatomic molecule HCl in the spectral region 3100-2600 cm^{-1} in dry nitrogen (NJ recorded at a resolution of 4 cm^{-1} and with a gas pressure of 400 torr under quasi-static condition. For a diatomic molecule like HCl which has no net spin or

orbital angular momentum (i.e., $^1\Sigma^+$ state) the spectrum contains only P and R branches according to the selection rules $\Delta J = \pm 1$. The assigned rotational quantum numbers, the wavenumbers for the P and R branches along with the absorbance is given in Table 4. The band origin v_0 of the 1-0 band is deduced from the measured wavenumbers of the P and R branches by plotting the values of R(J - 1)+ P(J) against J^2. The value of v_0 is found to be 2885.20 (5) cm^{-1} and is in agreement with the value of 2885.977 cm^{-1}(Bernath, 1995). The rotational constants B_1 & B_0 are also deduced in the present work and is found to be 10.092 (6) cm^{-1} and 10.401 (8) cm^{-1}. The temperature of the source of absorption is determine from the R(2) line of figure 5 and is found to be 187 K.

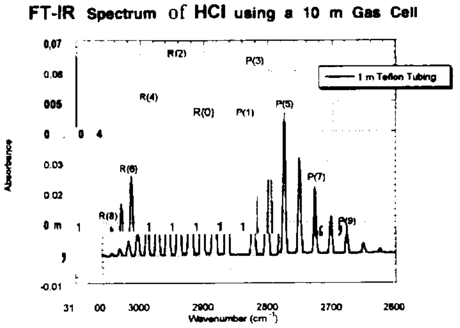


Figure 5
FT-IR spectrum of hydrogen chloride (HCl) in N_2 under quasi-static conditions at 4.0 cm-l resolution with a gas pressure of 400 torr

Table 4
FT-IR spectral peaks of hydrogen chloride (HCl) in N_2 under quasi-static conditions at a gas pressure of 400 torr using a 10 m Gas Cell

J	R(J) (cm^{-1})	Abs	P(J) (cm^{-1})	Abs
0	2905.63	0.0454	--------	---------
1	2925.18	0.0639	2864.27	0.0439
2	2944.10	0.0663	2842.78	0.0604
3	2962.37	0.0625	2820.90	0.0619
4	2980.08	0.0499	2798.18	0.0572
5	2997.13	0.0377	2775.02	0.0459
6	3013.70	0,0257	2751.42	0.0320
7	3029.36	0,0162	2727.02	0.0213
8	3044.40	0.0090	2677.03	0.0067
9	2058.57	0,0049	2651.44	0.0032
10	-------	--------	2624.95	0.0013

FT-IR Spectral Analysis of Methyl Nitrite

Figure 6 displays the FT-IR spectrum of 1% methyl nitrite (CH_3ONO) in helium (He) recorded at a resolution of 4 cm^{-1} and with a gas pressure of 300 torr. Table 5 identifies the chemical function groups and their corresponding wavenumbers for 1 % methyl nitrite.

The FT-IR spectrum of methyl nitrite exhibits various fictional groups. The functional group NH is exhibited at 3328.77 and 3219.93 cm-'. CH_3 occurs at 2956.90,2839.91, and 1388.86 cm-'. A double bond of nitrogen and oxygen is found at 1679.41 cm-'. A single bonded carbon-oxygen fictional group is observed at 1049.93 cm-'. The fictional group CH is located at 994.93 cm-'. A single bond of nitrogen and oxygen is found at 810.70 cm^{-1}. The functional group O-N=O is located at 625.49 and 565.80 cm-'.

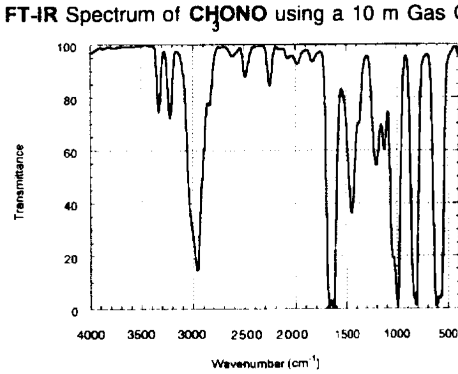


Figure 6
FT-IR spectrum of 1 % methyl nitrite (CH_3ONO) in He under quasi-static conditions at 4.0 cm-l resolution with a gas pressure of 300 torr

Table 5
FT-IR spectral peaks of 1 % methyl nitrite (CH_3ONO) in He under quasi-static conditions at a gas pressure of 300 torr using a 10 m Gas Cell

[illegible]	
3328.77 3219.93	NH
2956.90 2839.91	CH_3
1679.41	N=O
1388.86	CH_3
1049.93	C - O
994.93	CH
810.70	N-O
625.49	O-N=O

CONCLUSION

FTIR spectroscopy aided in the identification of multiple pollutants in the atmosphere. Owing to the reactivity of various sample materials with device materials used for QCM/SAW investigations, different tubings were used to check the acidity/basic nature of the sample gases in relation to various atmospheric trace gases. For example, both nitrogen dioxide (NO_2) and nitric acid (HNO_3) showed strong reactivity toward metallic materials like aluminum and copper.

The functional groups present in methyl nitrite (CH_3ONO) are important for the vibrational analysis of oxidation intermediates involved in the combustion of hydrocarbons and in air pollution. Such molecular radicals (e.g. methoxy) are also important in astrophysical study. Nitrogen dioxide, sulfur dioxide, and hydrogen chloride are trace gases that are present in the atmosphere and are important for the understanding of stratospheric ozone depletion, global warming, or/and photochemical smog formation (Sigrist, 1994).

FT-IR spectroscopy provides accurate values for the rotation constant of molecules (as determined for HCl in this paper) and the temperature of the infrared source of radiation used for recording the absorption spectra of molecule species.

REFERENCES

Abina, Rafiu A, Okabe H., & Misra, Prabhakar, FTIR Spectroscopy of HNO_3 and NO_2 Revelent to Stratospheric Wake Analysis, First NASA Student Conference, North Carolina A & T State University, 1996

Bernath, P. F., Spectra of Atoms and Molecules, New York, NY, 1995, Oxford University

Coleman, Patricia B., Practical Sampling Techniques for Infrared Analysis, Boca Raton, FL, 1993, CRC Press

Kamal, Mohammed, FT-IR Spectroscopy and High Resolution Laser Spectroscopy Associated with Alkoxy and Alkylthio Radicals (Master's Thesis), Howard University, Washington, DC, 1994

Sigrist, Markus W., Air Monitoring By Spectroscopic Techniques, New York, NY, 1994, John Wiley & Sons, Inc.

URC97069

-20 / 16

Cr^{2+} Diffusion Doping in ZnSe

Troy D. Journigan, K.-T. Chen, H. Chen, A, Burger
NASA/Fisk Center for Photonic Materials and Devices
Department of Physics, Fisk University
Nashville, Tennessee 37208
and
K. Schaffers, R.H. Page, and S.A. Payne
Advanced Solid State Laser Group
Lawrence Livermore National Laboratories
Livermore, CA 94550

Abstract

Chromium doped zinc selenides crystals have recently been demonstrated to be a promising material for near-IR room temperature tunable lasers which has an emission range of 2-3 pm. in this study, *a new* diffusion doping process has been developed for incorporation of Cr^{2+} ion into ZnSe wafers. This process has been successfully performed under isothermal conditions, at temperatures above 800 degrees Celsius. Concentrations in excess of 10^{19} Cr^{2+} ions/cm^3, an order of magnitude larger than previously reported in melt grown ZnSe material, have been obtained. The diffusivity was estimated to be about 10-* cm^2/sec using a classical diffusion model. Resistivity was derived from current-voltage measurements and in the 107- 10^{16} Ω-cm and increased as function of Cr concentration.

Introduction

Recently a new class of transition metal-doped (Cr, Ni, Co) Zn chalcogenides (Se, S, Te) materials have been investigated as potential candidates as tunable solid state lasers with a spectral emission range of 1-4 μm [1,2]. Such lasers can be employed by NASA as source for Light Detection and Ranging (LIDAR) systems. The LIDAR instrument transmits light out to a target, interacts with it and changed by the target. Some of the light is reflected/scattered back to the instrument where it is analyzed. LIDAR can be used as range finders, for remote sensing of chemical species (such as ozone, water vapor and pollutants) in the atmosphere and for remote measurement of the wind velocity. The absorption and emission spectroscopic properties have been measured to evaluate their potential for room temperature operation. Notably, the emission cross section and temperature-dependent lifetimes have been measured and they indicated that chromium ion has a potentially high gain cross section and low non-radioactive decay losses resulting in a high emission quantum yield. The quantum yield efficiencies are generally approaching 100% for chromium, in contrast to nearly 20% for cobalt, while nickel exhibited no room temperature luminescence. A pump volume of 10 x 6 x 3 mm and a Cr^{2+} concentration of 1.5 x 1 0^{20} ion/cm^3 are needed for a 1.63 pm (40 W, InGaAsP diode laser) pump light to be absorbed within about 100-200 pm from surface [1 ,2]. ZnSe has been previously doped with Cu, Ag, Ga, In, and Li for the achievement of suitable electrical properties. It has been reported that precipitation was observed for Ga[3], In[4], and Li[5] doped ZnSe. ZnSe has been doped with Cr^{2+} ion to investigate optical properties [6,7]. Based on the analysis of superhyperfine interactions with neighbors it has been reported that the chromium ion incorporated substitutionally into Zn lattice sites [8,9]. Up to date, there is little information about the usage of Cr doped ZnSe as solid state laser material and the doping process which can achieve Cr concentration as high as 10^{20} ion/cm. The motivation and scope of this work encompass the

innovative ideology of producing a high optical quality of ZnSe:Cr with Cr concentrations up to 10^{20} ion/cm^3 or as needed. It is the purpose of this paper to report our efforts at Fisk University with collaboration of the laser group at LLNL in the development of this process and evaluation of its usefulness in fabrication of a mid-IR solid-state laser.

Theory

Many of the thermally activated processes that occur in solids are diffusion controlled, and the rate of the process is determined by the diffusion rate of individual atoms from one atomic site to another. The driving force for atomic diffusion is the *concentration gradient* that exists between one point and another. In practice there are two particular cases of diffusion, steady state and nonsteady state diffusion, and the interdiffusion of ZnSe and chromium is a case of nonsteady state diffusion.

Consider the diffusion of solute atoms along the x direction between two parallel atomic planes perpendicular to the plane of the paper and separated by a distance a. Suppose that the planes are of unit area and that there are N solute atoms on plane (1) and N solute atoms on plane (2), where $N_1 > N_2$. Here N has the units number/area. The atomic concentration of solute atoms in number/volume on plane (1) is $c_1 = N_1/a$ and on plane (2) is $c_2 = N_2/a$. Thus, a concentration gradient $dc/dx = (c_2 - c_1)/a$ exists along the x direction. The *net flux* of diffusing solute atoms, J, from plane (1) to plane (2) is

$$J = -1/2a^2r(dc/dx) = -D(dc/dx) \qquad \text{(Eq.1)}$$

where the diffusion coefficient $D = 1/2a^2r$ and has units cm^2/sec. This equation is known as *Fick's first law of diffusion.* The minus sign indicates that atom flow occurs in the direction of negative concentration gradient.

Now we can relate the rate of accumulation of atoms to the change in concentration with time as

$$dc/dt = J_1 - J_2/\Delta x \text{ (atoms/cnf-see)} \qquad \text{(Eq. 2)}$$

As Ax becomes vanishingly small, it then reduces to

$$dc/dt = D(d^2c/dx^2) \qquad \text{(Eq. 3)}$$

which is known as *Fick's second law of diffusion.* In physical terms this law states that the *rate of compositional change* is proportional to the *rate of change of* the concentration gradient rather than to the concentration gradient itself. The solution of this latter equation depends on the boundary conditions imposed by the particular problem of interest. As an example, a semi-infinite slab (dimensions large compared with the diffusion distance) of initial composition c_0 has, for all $t > 0$, the concentration at the x = O interface maintained at a value c_1. This surface concentration is greater than the uniform initial composition of the slab, and consequently diffusion occurs into the slab (in the x direction) from the surface. The solution of the last equation mentioned and for the conditions described above is

$$c(x,t)v\text{-}c_0/c_1 - c_0 = 1 - \text{erf}\ (x/2\sqrt{Dt}) \qquad \text{(Eq. 4)}$$

where $c(x,t)$ is the concentration at some point x in the slab at $t > 0$ and erf is the *Gaussian error function*, as tabulated in mathematical tables. Examples of the type of concentration curve predicted by this equation are shown in Fig.B. A particularly interesting aspect of this equation, and equivalently of the solutions to numerous other diffusion problems, is that $c(x,t)$ is completely described at all x and t by the error function of $(x/2\sqrt{Dt})$. This has two important consequences. First, if we arc interested in a given composition, c', then

$$c' - c_0/c_1 - CO = 1 - \mathrm{erf}\,(x/2\sqrt{Dt}) \qquad \text{(Eq. 5)}$$

is constant, and therefore

$$x/\sqrt{Dt} = \text{constant} \qquad \text{(Eq. 6)}$$

We solve for x and *t so* that the concentration of the diffusing species is equal *to* one-half the value at the interface x = O. For these conditions

$$\mathrm{erf}\,(x/2\sqrt{Dt}) = 0.5$$

And from mathematical tables

$$x/2\sqrt{Dt} = 0.5$$

or

$$x = \sqrt{Dt} \qquad \text{(Eq. 7)}$$

Thus, as mentioned above, this simple relationship can be used as a rough estimate of time and temperature needed to achieve appreciable diffusion or redistribution of solute atoms over a distance x. And since many other solutions to Eq. (3) also indicate that Eq. (7) can be used to predict the time required for diffusion to occur over a given distance, it is not only a very simple but also a very useful relationship.

Experimental

The diffusion doping ampoules were made from fused quartz. The cleaning processes of ampoules includes two steps: (a) the ampoules were cleaned with aqua regia (25% HNO + 75% HCl) and rinsed with deionized water, (b) the ampoules were then heated at 100 °C above the diffusion temperature under 10^{-6} Torr vacuum. The Bridgman or seeded physical vapor transport (SPVT) grown ZnSe wafer, which were purchased from Eagle-Pitcher company, were loaded into precleaned ampoule along with CrSe powder (4N grade from Alfa company), The loaded ampoule was then sealed under higher range of 10-7 Torr vacuum. The sealed ampoule (typical length of 10 cm and 1.2 cm ID) was inserted into the middle section of a 12 inches long heat pipe which provided a isotherm environment in the middle 10 inches region between 400 and 1100 "C. and the heat pipe was then placed in art one zone furnace. A schematic drawing of the experimental arrangement is *shown in* figure 1. Several diffusion experiments have been performed under this doping process with temperatures ranging from 800 to 1000 °C for durations of either 2 or 5 day annealing periods.

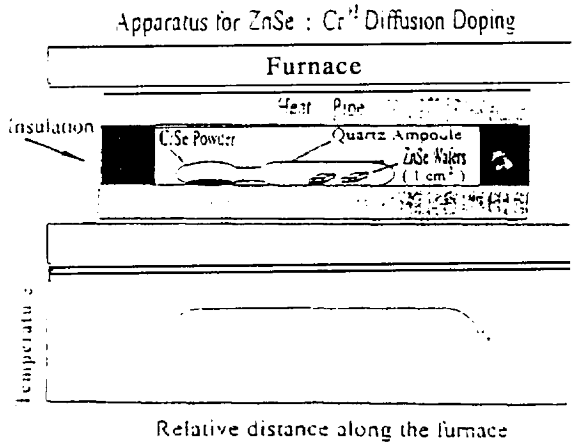


Figure 1, Experimental arrangement of diffusion doping process.

After annealing, the temperature was lowered by 5 °C/min to 400 °C, and then the furnace was shutdown and cooled to room temperature. The Cr^{2+} concentration, laser lifetime, and other optical properties were measured using optical absorption, inductively coupled plasma-mass spectroscopy, and a solid state laser testing by a solid state laser group in Lawrence Livermore National Laboratory [1,2]. Current-voltage measurement was performed at Fisk. Silver conducting paint and copper wire were used as metal contact and lead wire for connection to the I-V electronic system. A typical increment of 1 V and 1 min waiting time was used.

Results and Discussion

The following table lists the doping conditions, Cr^{2+} concentration, laser lifetime, and resistivity. The first column indicates the crystal growth method and its structural format. The first seven samples were grown by the Bridgman method and they are in the form of a polycrystalline (Poly.) and single crystalline structures with twin (Twinned).

ZnSe:Cr^{+2} Diffusion Experiment Results

Sample	Annealing Temperature (°C)	Annealing Time (days)	Concentration Cr^{+2} (cm^{-3})×10^{20}	Lifetime (μsec)	Abs. Coeff. Cr^{+2} (cm^{-1})	Abs. Coeff. Background (cm^{-1})	Resistivity (Ω-cm)
#12413: Bridg. Poly.	800	2	0.05	6	4.55	0.12	
#12413 A: Bridg Poly.	850	5	0.24	5	19.91	1.1	
#12429: Bridg. "twinned-top	900	5	0.19	5	16.41	2.84	
#12429A: Bridg Twinned-bot.	900	5	0.25		17.47	3.98	2.5×10^{7}
#12428; Bridg. Poly.	900	5	0.18	3	11.24	6.95	6.0×10^{11}
#12427: Bridg. Poly-top	950	2	0.38	2.5	26.64	4.72	2.0×10^{13}
#12426: Bridg. Poly- bot.	950	2	0.33	3.1	20.44	2.69	8.2×10^{11}
#12430: SPVT Single	950	1	0.17	6			9.9%10"
SPVT, Single	1000	5.5	1,02				1.9.10"

*All wafers were cut from ZnSe crystals, and purchased from Eagle-Pitcher Co.

The last two samples were single crystalline grown from seeded physical vapor transport method. The annealing temperature and Cr^{2+} concentration have been plotted in figure 2. For both the 2 and 5 day annealed samples, the Cr^{2+} concentration increases with annealing temperatures. This trend indicates that for 1000 "C and 5 days annealing conditions we have not reached saturation that would arise from the existence of a Cr^{2+} volubility limit in ZnSe. The linear least square fit lines (dashed line-2 days, solid line-5 days) seem to be fairly parallel to each other. The result proposed that the same Cr^{2+} concentration can be obtained by annealing at lower temperature for longer time as annealing at higher temperature for a shorter period of time.

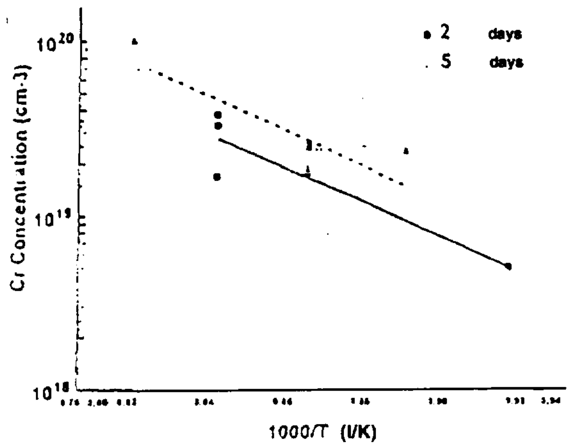


Figure 2. Plot of Cr^{+2} concentration under various annealing temperatures and times.

The possible disadvantages of high temperature annealing are the increase of the concentration of defects, such as vacancies and dislocations, and the deviation from stoichiometry of ZnSe. All samples have been polished after the diffusion doping process, and the color ranges from light reddish to darker reddish as the Cr^{2+} concentration increases. Due to the fact that the crystal growth temperature of SPVT grown ZnSe is around 1100 "C, it is a reasonable assumption that 1000 "C is very near the limit of annealing temperature for SPVT grown ZnSe wafer. The lifetime of single and polycrystalline wafer does not show a clear trend. This result maybe a good sign for industrial manufacture since that polycrystalline ZnSe has a much more lower production cost than single crystalline material. Figure 3 shows a typical I-V measurement result on the sample annealed at 1000 °C and 5 days.

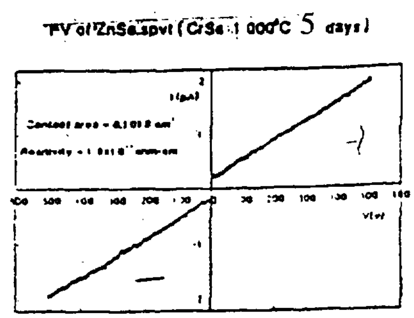


Figure 3.-4 typical I-V plot for sample annealed at 1000 °C and 5 days.

The resistance was derived from the slope of the I-V curve, and the resistivity was calculated by taking account of the contact area and thickness of the sample. The resistivity of Cr^{2+} doped ZnSe was found to be 5-8 times the order of magnitude higher than the previous undoped ZnSe. By assuming the equilibrium diffusion depth as 0.1 cm, which is deeper than the requirement to be used as solid state laser material, the diffusivity, D, can be estimated by a thin film diffusion model [10].

Summary

A new Cr^{2+} diffusion doped ZnSe process has been developed and demonstrated to yield Cr^{2+} concentrations as high as 1020 ions/cm^3 which is required for tunable solid state laser material use, The crystallinity of ZnSe wafers seem to have significant influence on the Cr^{2+} concentration in this doping process. The Cr^{2+} concentration has not reached its volubility limit in the ZnSe solid solution, and it could in the future be increased by annealing at temperatures above 10000 C or at a longer time and at lower temperatures. The resistivity of doped ZnSe increases dramatically by 5-8 orders of magnitude, in the range of 10^{13} - 10^{16} Ω-cm compared with undoped materials. The diffusivity of Cr in ZnSe has been estimated to be in the 10-8 cm^2/sec range.

Acknowledgments

Funding for this project have been supported by NASA through NASA/Fisk Center for Photonic Materials and Devices and by the DOE grant No. DE-F603-94 SF20368. The authors at Fisk University would like to acknowledge Dr. Kennedy J. Reed of Lawrence Livermore National Laboratory and Research Collaboration Program (RCP) for HBCU's *through* the LLNL Education Program Office.

References

[1] L.D. DeLoach, R.H. Page, G.D. Wilke, S.A. Payne, and W.F. Krupke, CALIOPE *Annual Meeting Technical Review Proceeding, 2,* July (1994).

[2] L.D. DeLoach, R.H. Page, G.D. Wilke, S.A. Payne, and W.F. Krupke, *Advanced Solid State Lasers Conference,* Memphis, Tennessee, January (1995).

[3] A.K. Ray and F.A. Kroger, *J. Electrochem. Sot., 125, 1355 (1978).*

[4] G. Jones and J. Woods, J. *Phys* D., 9, 799(1976).

[5] G.F. Neumark and S.P. Herko, *J. Cryst. Growth,* 59189 (1982).

[6] J.T. Vallin and G.D. Watkins, *Physical review B,* 9(3), 2051 (1974).

[7] G. Goetz, H. Zimmerman, and H.-J. Schulz, *Z. Phys. B,* **91,429 (1993).**

[8] M. DeWit, A.R. Reinberg, W.C. Holton, and T.L. Estle, *Bull. Am. Phys. Sot., 10,329* **(1 965).**

[9] T.L. Estle and W.C. Holton, *Phys. Rev.,* 150,159 (1966).

[10] Paul G. Shewmon, *"Diffusion in Solids,* " McGraw-Hill Book Company, London (1985).

[11] T. Lukaszewicz and J. Zmija, *J. Phys. Status Solid A, 62,696* (1980).

[12] M. Aven and R.E. Halstead, *Phys.* Rev., 137, A228 (1965).

URC97070

N 11 - 11

S70/02

Steady-State Solution of a Flexible Wing

Reza Karkehabadi[1]
Suresh Chandra[2]
Ramesh Krishnamurthy[3]
NASA Center for Aerospace Research
North Carolina A&T State University
Greensboro, NC 27411

A fluid-structure interaction code, ENSAERO, has been used to compute the aerodynamic loads on a swept-tapered wing. The code has the capability of using Euler or Navier-Stokes equations. Both options have been used and compared in the present paper. In the calculation of the steady-state solution, we are interested in knowing how the flexibility of the wing influences the lift coefficients. If the results of a flexible wing are not affected by the flexibility of the wing significantly, one could consider the wing to be rigid and reduce the problem from fluid-structure interaction to a fluid problem.

INTRODUCTION

With the advent of more powerful computers and more efficient algorithms, there has been significant advances in both computational fluid dynamics (CFD) and computational structural dynamics (CSD). These advances have played a major role and contributed significantly to the area of aeroelasticity. To increase the efficiency of aircrafts, particularly for high speed aircrafts, reducing the weight plays a major role and in doing so, this leads to a more flexible structure. Aeroelasticity has a major role in the design of aircrafts. The interaction of the flow with flexible structure could limit the performance of an aircraft and it also can cause dangerous situations. For example, due to the presence and movement of shock waves in the transonic range, undesirable aeroelastic behavior might occur. Also, a highly swept wing of an aircraft might experience vortex-induced aeroelastic oscillations [1].

There have been significant advances in the area of composite materials. Composite materials technology provides structural designers with a capability to specify many of the stiffness properties of modern aircraft structures. A structure can be designed in such a way that the deformation of the structure is prescribed by selecting the material and orienting the composite piles. This technology provides various possibilities for improving the aerodynamic performance of an aircraft under different loading conditions. To be able to solve fluid and structural equations simultaneously, helps in achieving these improvements.

Experimental tests, for aeroelastic wings, requires wind-tunnel experiments that are highly expensive. The experimental tests cannot be ruled out and it is necessary to be complemented with the numerical solutions. In order to calculate the aeroelastic response of a structure, the fluid and structural equations have to be solved simultaneously. NASA Ames Research Center has developed a code, ENSAERO, that is able to accurately couple the Euler and Navier-Stokes equations with the structural equations. This code calculates the aeroelastic response by simultaneously integrating the Euler/Navier-Stokes equations and structural equations of motion. The fluid equations are solved by using finite-difference technique and the structural equations are solved using finite-element method. An early version of ENSAERO [2] was applied to an elastic rectangular wing and the results demonstrated the accuracy of the code in predicting the flutter dynamic pressure of the wing. It should be noted that the Euler equation were used. ENSAERO was extended to solve Navier-Stokes equations [3]. The code has the ability to model moving control surfaces [4]. Also ENSAERO has the option of using Euler or Navier-Stokes equations and it is able to simulate transonic flows on wing-body combination [5]. Since grid generation techniques for aeroelastic calculations involve moving components, ENSAERO has the capability of using moving grids.

In the present paper, ENSAERO is used to compute the aerodynamic loads on a swept-tapered wing. The steady-state results are obtained by considering the wing to be elastic and the results are compared with the one obtained considering the wing to be rigid. Since we are interested in the steady-state solution and using a fluid-structure interaction code is computationally expensive. Is it possible to consider the wing to be rigid? Can we ignore the property of the material in obtaining the steady-state solution? We attempt to answer these questions.

---

[1] Research Associate
[2] Research Professor
[3] Research Associate

AERODYNAMIC EQUATIONS

The compressible N-S equations in Cartesian coordinates can be written as

$$\frac{\partial Q}{\partial t}+\frac{\partial E}{\partial x}+\frac{\partial F}{\partial y}+\frac{\partial G}{\partial z}=0 \qquad \text{(I)}$$

where Q, E, F, and G are flux vectors given by

$$Q=\begin{bmatrix}\rho\\ \rho u\\ \rho v\\ \rho w\\ E_t\end{bmatrix}$$

$$E=\begin{bmatrix}\rho u\\ \rho u^2+P-\tau_{xx}\\ \rho uv-\tau_{xy}\\ \rho uw-\tau_{xz}\\ (E_t+p)u-u\tau_{xx}-v\tau_{xy}-w\tau_{xz}+q_x\end{bmatrix}$$

$$F=\begin{bmatrix}\mathrm{P'}\\ \rho uv-\tau_{xy}\\ \rho v^2+p-\tau_{yy}\\ \rho uw-\tau_{yz}\\ (E_t+p)v-u\tau_{xy}-v\tau_{yy}-w\tau_{yz}+q_y\end{bmatrix}$$

$$G=\begin{bmatrix}\rho w\\ \rho uw-\tau_{xz}\\ \rho vw-\tau_{yz}\\ \rho w^2+p-\tau_{zz}\\ (E_t+p)w-u\tau_{xz}-v\tau_{yz}\ w\tau_{zz}+q_z\end{bmatrix}$$

and the shear stresses and heat-flux equations are given as

$$\tau_{xx}=\frac{2}{3}\mu(2\frac{\partial u}{\partial x}-\frac{\partial v}{\partial y}-\frac{\partial w}{\partial z})$$

$$\tau_{yy}=\frac{2}{3}\mu(2\frac{\partial v}{\partial y}-\frac{\partial u}{\partial x}-\frac{\partial w}{\partial z})$$

$$\tau_{zz}=\frac{2}{3}\mu(2\frac{\partial w}{\partial z}-\frac{\partial u}{\partial x}-\frac{\partial v}{\partial y})$$

$$\tau_{xy}=\tau_{yx}=\mu(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y})$$

$$\tau_{xz}=\tau_{zx}=\mu(\frac{\partial w}{\partial x}+\frac{\partial u}{\partial z})$$

$$\tau_{yz} \quad \tau_{zy} = \mu(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y})$$

$$q_x = -k\frac{\partial T}{\partial X},\ q_y = -k\frac{\partial T}{\partial Y},\ q_z = -k\frac{\partial T}{\partial Z}.$$

ρ is the fluid density and u, v, w are the fluid velocity in x, y, and z directions. Et is the total energy per unit volume, p is the pressure, μ is the viscosity, and k is the coefficient of thermal conductivity, The governing equation can be transformed "into general curvilinear coordinate where

$$\tau = t$$
$$\xi = \xi(x, y, z, t)$$
$$\eta = \eta(x, y, z, t)$$
$$\zeta = \zeta(x, y, z, t)$$

and the resulting transformed equations can be written in non-dimensional form as

$$\frac{\partial \hat{Q}}{\partial \tau} + \frac{\partial \hat{E}}{\partial \xi} + \frac{\partial \hat{F}}{\partial \eta} + \frac{\partial \hat{G}}{\partial \zeta} = 0 \qquad (2)$$

where ^ indicates the transformed quantities. In order to solve Navier-Stokes equations, extensive computational time and storage is required. For this reason, the reduced equations, known as "thin-layer" or "parabolized" Navier-Stokes equations are being used. In the thin-layer approximation the viscous terms containing derivatives in the parallel direction to the surface of the body are neglected from Navier-Stokes equations. The thin-layer model requires a boundary-layer type coordinate system. Equation 2 simplifies to

$$\partial_\tau \hat{Q} + \partial_\xi \hat{E} + \partial_\eta \hat{F} + \partial_\zeta \hat{G} = \mathrm{Re}^{-1}\, \partial_\zeta \hat{S}$$

where all the viscous terms are contained in S. In order to solve the aerodynamic equations, ENSAERO has capability of both central-difference and upwind scheme [6]. The central difference scheme used is based on the implicit factorization algorithm of Beam and Warming [7] with the modifications by Pulliam and Chaussee [8] for diagonalization.

AEROELASTIC EQUATIONS OF MOTION

The aeroelastic equations of motion for a flexible wing are solved by using the Rayleigh-Ritz method, Using this technique, the aeroelastic displacements at any time are expressed as a function of assumed modes. The number of assumed modes are Finite and the amount of contribution that each mode has on the total motion is derived by Lagrange's equation. For further detail and more information see [2]-[6].

The finite element matrix form of the aeroelastic equations of motion for the elastic wing are given in general form as

$$[M]\{\ddot{q}\} + [G]\{\dot{q}\} + [K]\{q\} = \{F\}$$

where [M], [G], and [K] are the mass, damping, and stiffness matrices, respectively. { q } is the displacement vector and { F } *is* the aerodynamic loading vector and is computed by solving Euler or Navier-Stokes equation. Dots denote the derivative with respect to time.

RESULTS

A swept-tapered wing, shown in Figure 1, is used in our calculations. The aspect ratio of the wing is 3, the tapered ratio is 0.14, and the leading edge sweep angle is 51.34 degrees. The Mach number is 2.5 and Reynolds number is 5×10^7. As a first step, we consider the wing to be elastic and solve the fluid and structure equations simultaneously by using ENSAERO. Total lift coefficient for this wing as a function of time is shown in Figure 2. The angle of attack is 10 degrees. The figure compares the results obtained using Euler with Navier-Stokes equations. As the figure indicates, the difference in total lift between these two options is not significant.

Figures 3, 4, and 5 compare the sectional lift obtained using Navier-Stokes and Euler equations. The wing is set into motion from the impulsive start until it reaches steady-state. The sectional lift coefficients are shown for different sections on the wing, 28%, *61%,* and 90% from the root of the wing. The lift coefficients oscillate for all

cases because the wing is elastic and, in our calculations, we consider the elasticity of the wing. Again, there is not a significant difference in the sectional lift coefficients between Euler and Navier-Stokes equations. As one would expect, the lift coefficients of the section that is closer to the root reaches steady-state faster than the other sections. This is more clear by comparing Figures 3,4, and 5 where the lift coefficients are shown for the different sections of the wing. At time= 94.3 seconds the section 28% away from the root has little oscillation and the one that is 90°/0 away from the root has the most.

In order to obtain the steady-state solution for an elastic wing, we have considered the flexibility of the wing in our calculations. This requires solving fluid and structural equations simultaneously. Calculating the aerodynamic loads on an elastic wing requires more computation and therefore, it is computationally expensive and time consuming. In calculating the steady-sate solution, often the wing is considered to be rigid. We consider the above wing to be rigid and calculate the aerodynamic loads. Again, the angle of attack is 10 degrees. The total lift coefficient on the wing is shown in Figure 6. The solid line is the Navier-Stokes solution and the dashed line is the Euler solution. Comparing the lift coefficients of the rigid wing (Figure 6) with the elastic wing (Figure 2) shows that the steady-state results differ significantly and the difference is about 19%. The sectional lift coefficient for the rigid wing at 28°/0 from the root is shown in Figure 7. The difference with the elastic case is about 13°/0. In Figure 8 the sectional lift coefficient is shown for the section of the wing located at 6 1% from the root. Comparing this figure with Figure 4, one sees that the steady-state lift coefficient is 0.165 for rigid and 0.128 for elastic case. The difference between the rigid and elastic is about 29%.

There is more flexibility near the tip; hence, the difference with the rigid case is highest near the tip. In Figure 9 the sectional lift coefficient is shown for the wing at 90% from the root, which is very close to the tip. Again, the wing is considered to be rigid. Figure 5 demonstrates the sectional lift coefficient for the same section with the same condition except the wing is elastic. The difference between these two are significantly large, 60%. In all cases, the Navier-Stokes solution is given as well as Euler solution. The Figures indicate that the lift coefficients using either options are not being significantly affected.

CONCLUSION

It was shown that the computed values of lift coefficients are not greatly affected, for the configuration we have considered, by using either Navier-Stokes or Euler equations. The elasticity of the wing contributes to the lift coefficients. A more rigid wing has higher lift coefficients. In case of a flexible wing, part of the energy is used to oscillate the wing; hence, the value of the lift decreases with the increase in flexibility of the wing. The section" of the wing that is near the tip has the highest amount of flexibility; therefore, the difference (with the rigid one) is highest in this section. In order to obtain the lift coefficients on a wing one has to consider the elasticity of the wing and in reducing the problem to a rigid wing some accuracy might be lost.

ACKNOWLEDGMENTS

The authors would like to thank NCSC for the computer time that they have provided. Our thanks are extended to NASA for their support under the grant NAGW-2924.

REFERENCES

1. Dobbs, S. K., and Miller, G. D., "Self-Induced Oscillation Wind-Tunnel Test of a Variable Sweep Wing," AIAA Paper 85-0739, April 1985.
2. Guruswamy, G. P., "Unsteady Aerodynamic and Aeroelastic Calculations for Wings Using Euler Equations," AIAA Journal, Vol. 28, 1990, PP. 461-469.
3. Guruswamy, G. P., "Navier-Stokes Computations on Swept-Tapered Wings, Including Flexibility," AIAA Paper 90-1152, Apr. 1990.
4. Obayashi, S., and Guruswamy, G. P., "Navier-Stokes Computations for Oscillating Control Surface," AIAA Paper 92-4431, Aug. 1992.
5. Obayashi, S., and Guruswamy, G. P., and Tu, E. L., "Unsteady Navier-Stokes Computations on a Wing-Body Configuration in Ramp Motions," AIAA Paper 9 1-2865, Aug. 1991.
6. Obayashi, S., and Guruswamy, G. P., and Goorjian, P. M., "Streamwise Upwind Algorithm for Computing Unsteady Transonic Flows Past Oscillating Wings," AIAA Journal, Vol. 29, No. 10, Oct. 1991, PP 1668-1677.
7. Pulliam, T. H., and Chaussee, D. S., "A Diagonal Form of an Implicit Approximate Factorization Algorithm," Journal of Computational Physics, Vol. 39, No. 2, Feb. 1981, pp 347-363.
8. Beam and Warming, R. F., "An Implicit Finite-Difference Algorithm for Hyperbolic Systems in Conservation Law Form," Journal of Computational Physics, Vol. 22, No. 9, Sept. 1976, pp. 87-110.

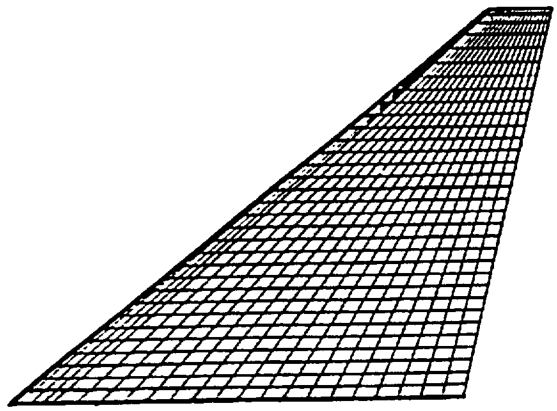


Fig. 1: 2-D view of the wing used in this paper.

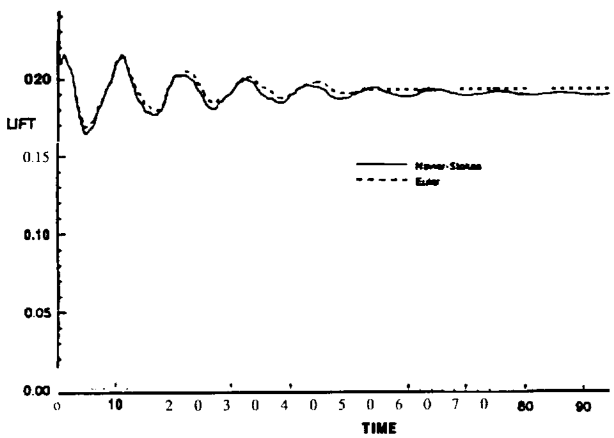


Fig. 2: Total lift coefficients of a flexible wing as a fiction of time.

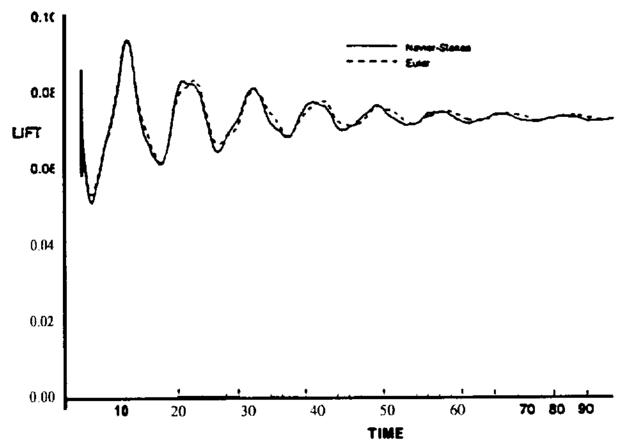


Fig. 3: Lift coefficients of a flexible wing as a function of time for the section 28% from the root.

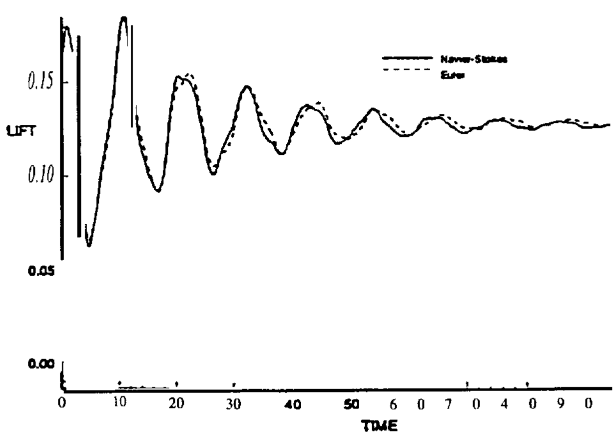


Fig. 4: Lift coefficients of a flexible wing as a function of time for the section 61'%. from the root.

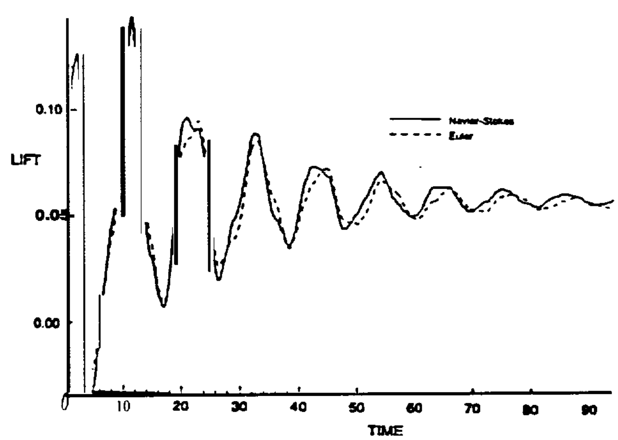


Fig. 5: Lift coefficients of a flexible wing as a function of time for the section 90% from the root.

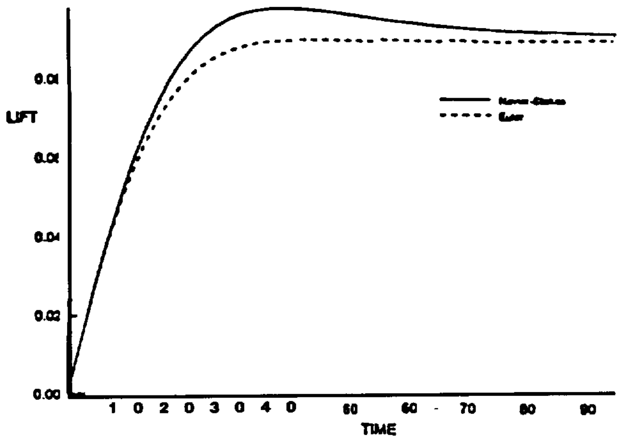


Fig. 6: Total lift coefficients of a rigid wing as a function of time.

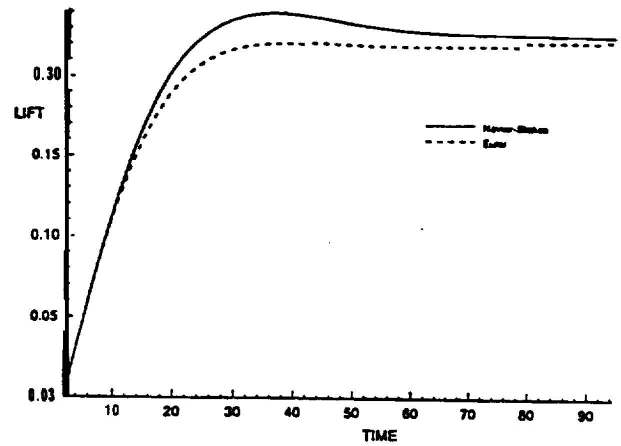


Fig. 7: Lift coefficients of a rigid wing as a function of time for the section 28% from the root.

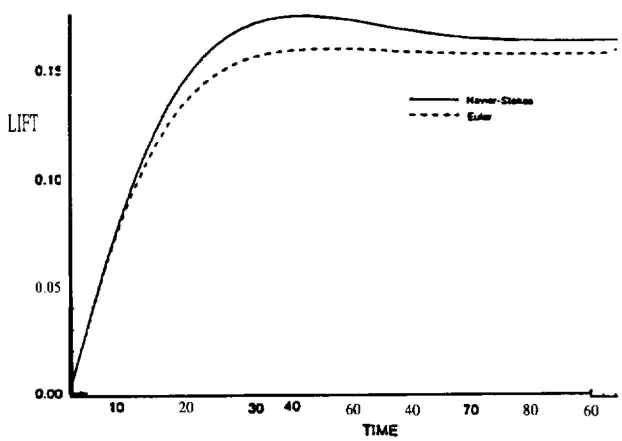


Fig. 8: Lift coefficients of a rigid wing as a function of time for the section 61% from the root.

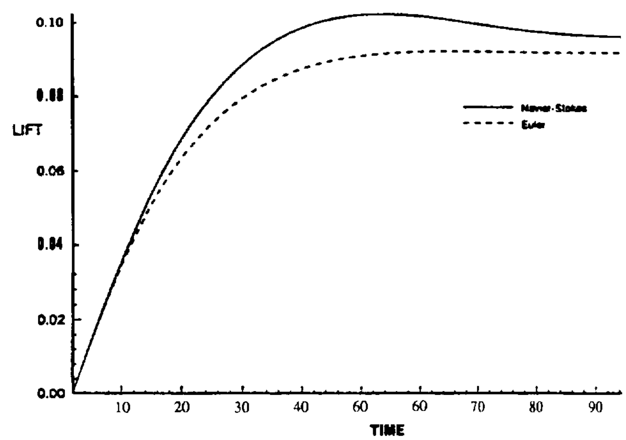


Fig. 9: Lift coefficients of a rigid wing as a function of time for the section 90% from the root.

11/25

URC97071

A new probe to change Curie temperature of $PbTiO_3$ sensors

R. S. Katiyar and Meng Jinfang *

Department of Physics, University of Puerto Rico, San Juan, Puerto Rico 00931-3343, USA

High temperature Raman spectra of nanocrystalline $Pb_{0.8}Ba_{0.2}TiO_3$, $Pb_{0.8}Sr_{0.2}TiO_3$, $Pb_{0.8}La_{0.2}TiO_3$ and $PbTi_{0.75}Zr_{0.25}TiO_3$, have been measured, as a function of particle size. There appears respectively a distinct temperature-induced soft mode phase transition in every sample whose Curie temperature can be determined from the mean-field theory. The detailed Curie temperature shift in modified $PbTiO_3$ ceramics by Ba, Sr, La and Zr, has also been investigated as a function of particle size, This study will favor preparations of high efficiency $PbTiO_3$ sensors with an adjustable Curie temperature.

I. INTRODUCTION

It is well known that pyroelectric infrared sensors show good sensing characteristics. Differing from photon sensors, they can be used at room temperature and their IR response does not depend on the wavelength of the IR radiation used. By combining these promising pyroelectric effects with sophisticated silicon IC technology, a realization of silicon monolithic IR sensor at near ambient temperature could be expected for a wide variety of applications. Nowadays, various kinds of pyroelectric materials like $LiTaO_3$ or TGS single crystals are commercially available. Among these materials, $PbTiO_3$ is a good candidate for IR detector because it has a large spontaneous polarization P_s(75μC/cm^2) and a small dielectric constant (~ 100) along the polarization axis [1]. $PbTiO_3$ is a perovskite type of ferroelectric isomorphous with $BaTiO_3$ and its tetragonality(c/a=1 .06) is also very large 2. But the Curie temperature T_c of $PbTiO_3$ single crystal is quite high, around 490 °C. For better sensitivity of detectors, the Curie temperature must be close to the operating temperature, In this sense, if the Curie temperature of $PbTiO_3$ could be reduced to about 160 °C by ionic substitution, its pyroelectric efficiency at room temperature can be strongly influenced. Obvious] y, the decrease of the Curie temperature T_c of $PbTiO_3$ can be realized by the substitution of of Pb or Ti ions. However, the substituted composition must have a suitable quantity in order to have appropriate T_c with high property coefficients. The experimental results show that the part substitution of Ba for Pb will lead to a decrease in Q and $\varepsilon_{33}/\varepsilon_0$, but an increase in k_p and a good frequency-temperature property. Practically, for materials used as ceramic wave filters and sonar emission vibration pickup, a high Q will be needed, But, in general, the material with a high Q has a low k_p. In materials with high ε there appears a large tanδ and the materials with high k_p will probably lead to an unstable physical property. This reveals that it is very difficult to get an idea sensor with an appropriate T_c and high property coefficients only by dopants. Thus, it is essential to probe other ways to resolve this contradiction. So far, effects of particle size on physical properties and the Curie temperature T_c of materials have been widely received an increasing interest and attention[3-9]. For example, investigations for $PbTiO_3$ show that its Curie temperature can be lowered to room temperature as particle size is decreased to 12.6-20 nm. In this paper, we have made a detailed study for the modified $PbTiO_3$ by Ba, Sr, La and Zr as a function of particle size using Raman scattering spectroscopy.

II. EXPERIMENTAL

Nanocrystalline $Pb_{0.8}Ba_{0.2}TiO_3$(PBT20), $Pb_{0.8}Sr_{0.2}TiO_3$,(PST20), $Pb_{0.8}La_{0.2}TiO_3$ (PLT20) and $PbTi_{0.75}Zr_{0.25}TiO_3$(PZT25) with different particle sizes were separately prepared, using a sol-

gel process in which barium acetate, lanthanum acetate, strontium acetate, lead acetate, zirconium and titanium butoxide were used as the precursor materials. The processes involved dissolving the metal-containing compounds in the solvent, hydrolyzing and polycondensing the resulting solution into various gels, and finally heat treating these gels at different temperatures and for different time intervals to form nanocrystalline powders. Chemical phase analysis was done using powder XRD measurements. The average grain size was calculated from the full width at half maximum(FWHM) of the (111) diffraction peak using the Scherrer equation $d_{XRD}=K\lambda/\beta(\theta)\cos\theta$ [10]. Where λ is the x-ray wavelength, $\beta(\theta)$ is the FWHM of the diffraction line, θ is the angle of diffraction, and the constant $K\approx 1$. Raman spectra of the samples were measured using a Spex- 1403 Raman spectrometer with a double monochromator, a standard photon-counting technique and an Ar^+ ion laser [1].

III. RESULTS AND DISCUSIONS

Raman spectra of PBT20, PST20, PLT20 and PZT25 with particle sizes of 60 nm, 37 nm, 54 nm and 44 nm are respectively shown in Fig. 1a - d. These spectra reveal the tetragonal structure of nanocrystalline materials at room temperature, similar to one observed for bulk $PbTiO_3$(PT) [12]. In earlier work, Ishikawa et al. had investigated samples with different average sizes as a function of temperature [13]. The T_c was indirectly measured as the temperature at which the frequency (ω_s) of the soft E(TO) mode ($\omega_s \to 0$ as $T \to T_c$) vanished. T_c was found to decrease with decreasing particle size with d_{crit}= 12.6 nm. On the other hand, in the case of bulk materials, the first phase transformation in $PbZr_xTi_{1-x}O_3$ (PZT) system was found to occur at x=0.52 [14-15], whereas no phase transformation was observed in bulk $Pb_{1-x}Sr_xTiO_3$ (PST) and $Pb_{1-x}Ba_xTiO_3$ (PBT) systems as x is increased from 0.0 to 1.0 16-17. Recent investigations on nanocrystalline PST, PBT, PZT and PLT with particle sizes of 37nm, 60 nm, 44 nm and 54 nm have shown that a phase transformation occurs at x=0.7, 0.7, 0.4 and 0.4 respectively [18-20]. According to these facts mentioned above, the samples used in present experiment are confirmed to have a tetragonal ferroelectric structure at ambient temperature.

From Fig. 1, it is found that the low frequency phonon mode of every sample, the E(TO) soft mode, shows a decrease in frequency, i.e. "softening", and widening in linewidth, on increasing temperature, as observed in pure bulk $PbTiO_3$. The dependence of the squared frequency of the soft mode upon temperature for PBT20, PST20, PLT20 and PZT25 is respectively displayed in Fig. 2a - d, revealing a nearly linear temperature dependence, in accordance with the soft mode theory in the mean-field approximation 21.

It should be noted that, because of a strong stray light in these materials, especially in PBT20, it is very difficult to measure the lowest frequency of the soft mode close to T_c. This is merely done by extrapolation of the observed data. The Curie temperature of pure $PbTiO_3$ is found to decrease on decreasing the particle size. Thus, it is reasonably considered that PBT20, PST20, PLT20 and PZT25 also undergo a similar temperature-induced soft mode as in pure $PbTiO_3$.

A detailed particle size dependence of the Curie temperature in the above mentioned materials has been performed and illustrated in Fig. 3a-d. One may notice that the transition temperature from cubic (CP) to tetragonal (TP) is shifted towards a lower temperature with decreasing particle size in nanocrystallines.

A careful observation reveals that the particle size corresponding to a Curie temperature of 20 ^oC in PBT20 and PST20 is smaller than that in PLT20 and PZT25. The particle size dependence of Curie temperature in PZT25 is also found to have a slow change above the particle size of 40 nm, which differs from that in PBT20, PST20 and PLT20. This may originate from many factors, for example, different chemical properties of Ba, Sr, La and Zr as well as their different ion radii(Ba^{2+} -0.143 nm, Sr^{2+} -0.127 nm, La^{3+} ~ 0.122 nm, Zr^{4+} -0.079 nm) will lead to various distortions of TiO_6 octahedra thereby changing the short-range forces responsible for the phase transitions. In addition, the particle size dependence of T_c may result from different substitution forms of Ba, Sr and La ions for Pb ions at A sites, and Zr for Ti ions at B sites in ABO_3 perovskite oxides.

The softening and widening of the soft mode caused by decreasing particle size, as observed by increasing pressure, have been measured in many nanocrystalline materials including $PbTiO_3$.[13] As particle size of materials is decreased, the surface stress on particles will produce an additional effect on atoms inside particles due to some interfaces and spherical particle distribution. The additional effect is similar to that of hydrostatic pressure on atoms. Hence, the decrease of the Curie temperature with reducing particle size results from effects of surface stress of spherical particles.

IV. CONCLUSIONS

High temperature Raman spectra for PBT20, PST20, PLT20 and PZT25 with various particle sizes have been measured and their Curie temperature can be determined by the temperature dependence of the squared frequency of the soft mode which is consistent with the mean field approximation. The Curie temperature is found to shift towards a lower temperature with decreasing particle size and it depends upon the substitution forms of doping ions for the ions at A and B sites in ABO_3 perovskite materials. The surface stress from small particles plays an important role in lowering Curie temperatures relative to the corresponding bulk materials. Our study will offer a physical basis for preparing $PbTiO_3$ sensors with an adjustable Curie temperature.

* On leave Department of Physics, Henan University, Kaifeng, P. R. China

ACKNOWLEDGMENTS

The work was supported by the NASA-NCCW-0088 and NSF-OSR-9452893.

REFERENCE

1. V. G. Gaurilyachenko, R. I. Spinko, et al., *Sov. Phys.* Solid *State* 12, 1203(1970).
2. G. Shirane, R. Pepinsky and B. C. Frazer, *Act Crystallogr.* 9, 13 1(1956).
3. K. K. Deb, *Ferroelectrics, 82, 45(1988).*
4. Soma Chattopadhyay, Pushan Ayyub, V. R. Palkar et al., *Phys. Rev.* **B52**, *13 177(1995).*
5. Jinfanf Meng, Guangtian Zou et al., *J. Phys.: Condens Matter, 6, 6543(1994).*
6. R. Bachmann and K. Baerner, ***Solid State Commun. 68(9),865(1988).***
7. K. Uchino, E. Sadanaga, T. Oonish, T. Morohashi, H. Yamamura, ***Transactions. 8, 107(1990).***
8. T. Yamamoto, K. Urabe and H. Banno, ***Japanese Journal of Applied Physics, 32, 4272(1993).***
9. Stefan Schlag and Hans-Friedrich Eicke, ***Solid State Commun. 91, 883(1994).***
10. A. J. C. Warren, Proc. Phys. Sot. London 80, 286(1962).
11. Guangtian Zou, Jinfang Meng et al., *Phys. Lett. A,* ***175, 246(1993).***
12. R. J. Nelmes and W. F. Kuhs, ***Solid State Commun. 54, 721(1985).***
13. K. Ishkawa, K. Yoshikawa and N. Okada, ***Phys. Rev.* B37**, 5852(1988).
14. Wenwu Cao and L. Eric Cross, ***Phys. Rev. B, 47(9), 4825(1993).***
15. B. Jaffe, W. R. Cook and H. Jaffe, Piezoelectric Ceramics. Academic Press, New York (1970).
16. G. Burns and F. H. Dacol, *J. Raman Spectrosc.* 10, 227(1981).
17. G. Burns and B. A. Scott, ***Solid State Commun. 9, 813(1971).***
18. Jinfang Meng, Guangtian Zou et al. J. *Phys.:* ***Condense Matter, 6, 6549(1994)***
19. Jinfang Meng, Guangtian Zou and Qiliang Cui, ***Since in China, 25, 420(1995).***
20. Jinfang Meng, K. Rai. Brajesh and R. S. Katiyar, to be publidhed.
21. M. E. Lines and A. M. Glass, ***Principles and Applications of Ferroelectrics and Related Materials, 1979.***

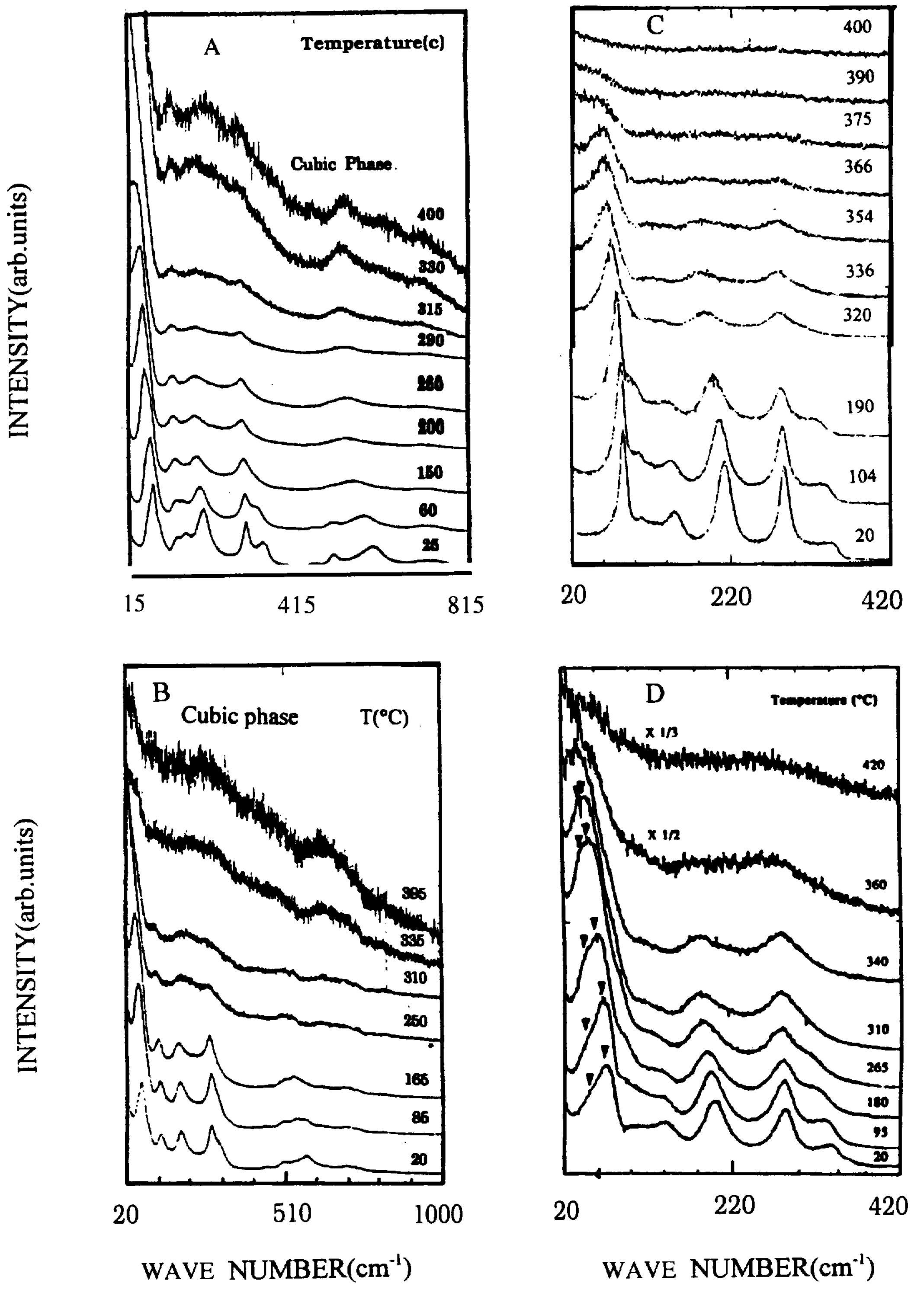


Fig. 1 Raman spectra of nanocrystalline materials at various temperatures
A. $Pb_{0.8}Ba_{0.2}TiO_3$ B. $Pb_{0.8}Sr_{0.2}TiO_3$
C. $Pb_{0.8}La_{0.2}TiO_3$ D. $PbTi_{0.75}Zr_{0.25}O_3$

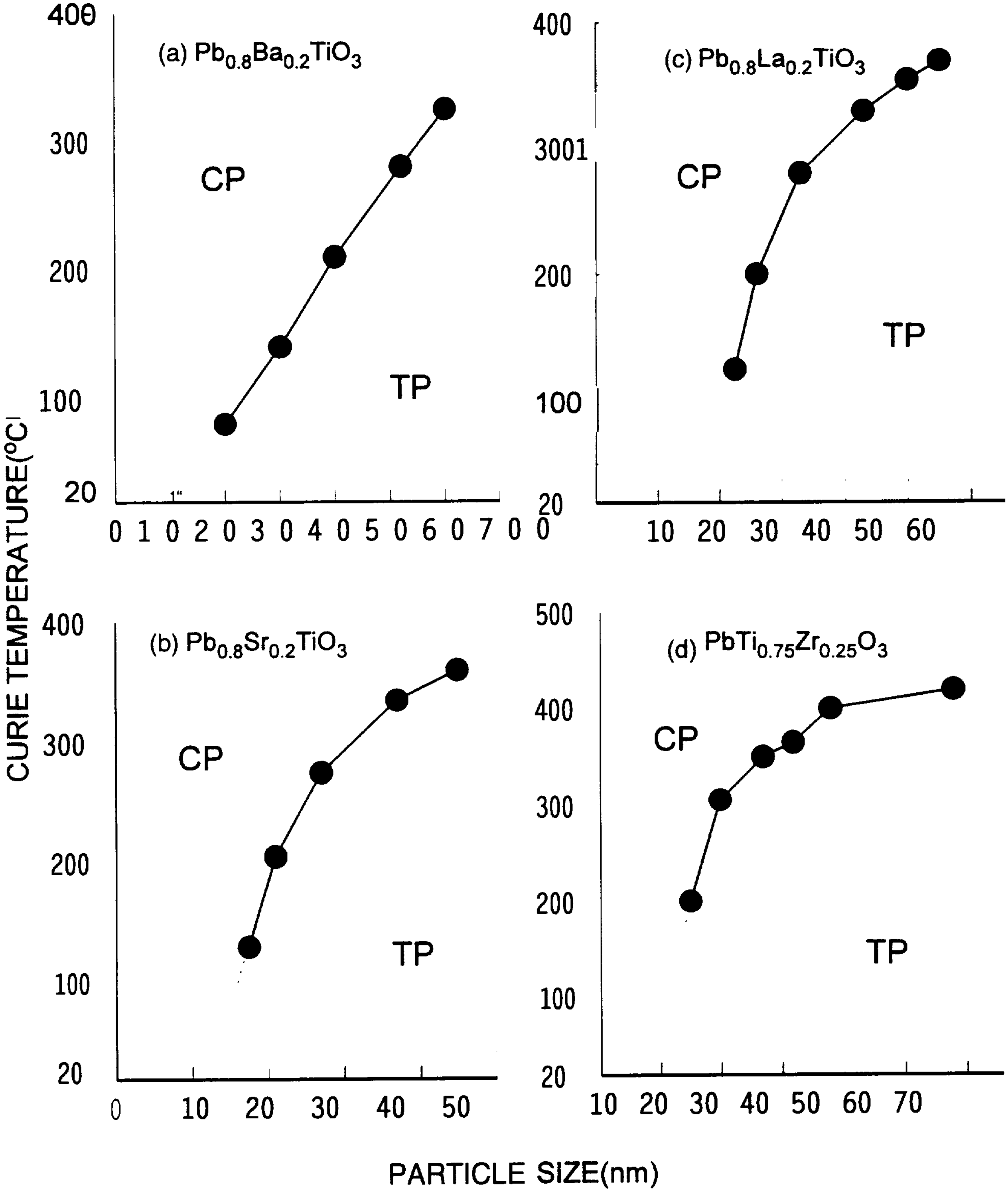


Fig. 3 The particle size dependence of Curie temperature for nanocrystalline

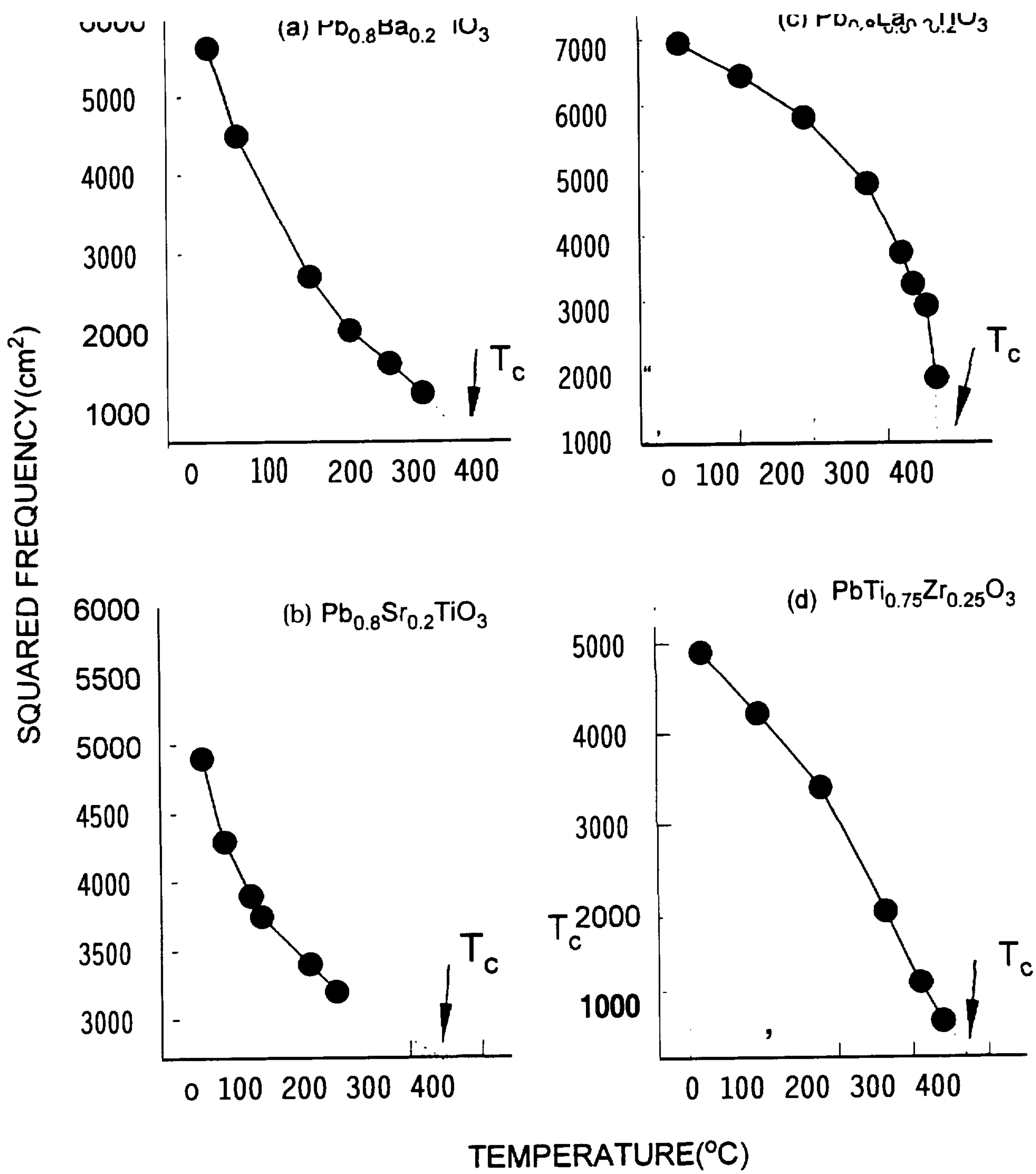


Fig. 2 The temperature dependence of the squared frequency of the soft mode

Page intentionally left blank

393 / 46

URC97072

An Integrated Remote Sensing and Geophysical Analysis of the Upper Crust of the Southern Kenya Rift

G. Randy Keller and Silas Simiyu

Pan American Center for Earth and Environmental Studies
University of Texas at El Paso
El Paso, Texas 79968

INTRODUCTION

The southern Kenya Rift (Figure 1) has been the target of a number of recent geophysical and geological investigations. This is due in part to the Kenya Rift International Seismic Project (KRISP) which has focussed attention on the region, but is also the result of the Kenya Power Company's geothermal exploration efforts. These studies have created a large new data base which has not been fully exploited to study upper crustal structure. The results presented here are the result of the combined analysis of satellite imagery, new gravity measurements, drillhole data from geothermal areas, KRISP seismic data, and geologic constraints to conduct an integrated analysis of upper crustal structure. Our focus on the upper crust is different from the previous studies motivated by the KRISP effort which have targeted lithospheric scale structure. The combination of all of these data allows us to pursue a level of detail which has not been possible previously. Our approach included simultaneous 2–D modeling of first arrivals from KRISP seismic lines in the area (Figure 1) and $2^1/_2$–D modeling (Cady, 1980) of gravity profiles along these lines, modeling of gravity profiles across key features where drill holes (Simiyu, 1994; Mwaura, 1990) provided important constraints, and construction of satellite images and gravity maps to provide a view of the lateral extent of features revealed in the modeling. Throughout this process, whatever supporting data which were available were employed as constraints. The models produced should be considered as geologic cross sections constructed with the aid of geophysical data not as geophysical models whose details can be defended on the basis of a single data set.

The classic rift valley in southern Kenya is a 50-70 km wide volcanic filled depression bounded by major normal faults and is largely located within the Proterozoic Mozambique mobile belt. Rifting is thought to have been initiated above a basement shear system that marks the contact between the Archean craton and the mobile belt (e.g., Smith, 1994; Smith and Mosley, 1993).

The latest volcanic activity within the rift valley occurred in the axial region and resulted in the formation of caldera volcanoes and volcanic cones (Figure 1). Menengai,

which is the most northern volcano within the Nakuru–Naivasha basin, is located about 60 km north of Eburru and consists almost exclusively of strongly peralkaline oversaturated trachytes. Eburru, is dominated on the surface by pantellerites, and pantelleritic trachytes, but trachytes are more abundant with depth. Syenite intrusive were encountered at the bottom of some of the holes drilled in the volcanic complex for geothermal exploration, The Elmenteita and Ndabibi basaltic fields are located northeast and south of Suswa (Figure 1) respectively and both lie on the N–S axial fault system.

GRAVITY ANALYSIS

Analysis of the gravity field shows that the regional negative anomaly over the Kenya rift between 10 N and 3° S is approximately two dimensional striking NS with its major axis displaced slightly to the west of the rift valley. There is also a NW–SE trend in the central part of the study area south of latitude 10 S. The traditional Bouguer anomaly map of the region is dominated by the regional gravity low due to deep structures which are not of interest in this study.

In order to better define shallow features of interest in this study, a variety of bandpass filters were applied to the Bouguer anomaly values. The maps for wavelengths of 30–1 50 km and 50–1 500 km were chosen for further analysis. The 30–150 km bandpass map does not include long wavelength (>150 km) features associated with the deep mantle anomalies of the Kenya dome. Wavelengths less than 30 km were also removed since such short wavelength anomalies are caused by very shallow sub-surface inhomogeneities beneath individual gravity stations. We feel that this map best depicts the features targeted in this study. In particular, the axial gravity high is very well defined as a narrow (20 ± 5km) gravity high extending from Lake Baringo to Lake Naivasha. South of Lake Naivasha, this anomaly widens to a maximum of 40 km at Suswa and dies 30 km south of the Suswa volcano, At Lake Magadi and further south, the rift valley occurs within a broad gravity low which may indicate either that more low-density fill is present, the rift beneath Magadi developed on a different crustal type, or the crust at Lake Magadi has not been magmatically affected in the same way as in the northern sector.

Our next step was to extend the seismic results employing gravity data for which the geographic coverage was extensive. The gravity modeling employed a $2^1/_2$–D forward modeling scheme. In this study, densities and depths to particular interfaces were constrained using seismic and drilling data to maximum extent possible. Five gravity profiles were constructed across the apex of the Kenya dome. The location of these profiles was governed by data coverage and location of seismic and drill hole constraints.

DISCUSSION AND CONCLUSIONS

Results from this study show that the southern Kenya rift has a complex crustal structure consistent with the complex tectonic history that the region has undergone. Because of the deep drill holes associated with geothermal exploration, the best constrained cross-section can be drawn through the Olkaria region (Figure 2). In this area, Bosworth (1989) attributed the sinuous course of the rift to the en-echelon arrangement of faults defining a series of asymmetric half grabens (basins) where the half graben change polarity at accommodation zones. In his model, extension is accommodated along major listric detachments of lithospheric extent which alternate in polarity along the length of the rift. They suggested that the major detachment fault in the Nakuru–Naivasha sub-basin is the Sattima fault which lies on the eastern margin at the Aberdare Range. This study suggests a different model since we show the rift graben to be deepest in front of the western margin at the Mau escarpment as opposed to the eastern rift margin. This rift basin fill geometry is clearly shown on the modeled gravity profiles across the Naivasha and Suswa areas and the seismic model north of Suswa volcano. The overall implication of this result is that the southern and central Kenya rift valley contains three large, asymmetric rift basins with east directed polarity. This result is different from the Western branch of the East African rift system where polarity reversals are well documented (Rosendahl, 1987). We suggest in this study that uniform asymmetry with east-dipping detachment faults in the southern and central Kenya rift is related to inherited mechanical anisotropies in the basement. It is therefore consistent to conclude from this study that there is an intimate relationship between rift trend, detachment geometry and pre–existing lithospheric grain.

Simultaneous visualization of the satellite and gravity data shows the presence of a series of gravity highs coincident with the locations of large volcanic centers. Apart from the intrusions at these volcanic centers, the entire upper crust of the southern Kenya rift shows evidence of magmatic injection by dike swarm activity (Swain, 1992) because seismic velocities and densities are higher than in adjacent areas. Thus, the crust has been extensively modified both by dike injection an intrusion of discrete magma bodies.

REFERENCES

Bosworth, W., 1987. Off-axis volcanism in the Gregory rift, East Africa: Implications for models of continental rifting, Geology, 15:397400.

Cady, W. J., 1980. Calculation of gravity and magnetic anomalies of finite–length right polygonal prisms. Geophysics, 45: 1507–1 512

Mwaura, M. N., 1990. Boreholes drilled for groundwater in the southern rift, Kenya. Geology, maintenance, productivity and water quality. Report to the Ministry of Water Development, Nairobi, Kenya. 126pp.

Rosendahl, B. R., 1987. Architecture of continental rifts with special reference to East Africa. Ann. Rev. Earth Planet. Sci., 15, 445–503.

Simiyu, S.M. 1994. Gravity interpretation of the Suswa and Domes geothermal prospects. A KPC review report.

Smith, M., 1994. Stratigraphic and structural constraints on mechanisms of active rifting in the Gregory Rift, Kenya. In: C. Prodehl, G.R. Keller and M.A. Khan (Editors), Crustal and Upper Mantle Structure of the Kenya Rift. Tectonophysics, 236: 3–22.

Smith, M. and Mosley, P. N., 1993. Crustal heterogeneity and basement influence on the development of the Kenya Rift, East Africa. Tectonics, 12:591-606.

Swain, C. J., 1992. The Kenya Rift axial gravity high: a re–interpretation. Tectonophysics, 204:59–70.

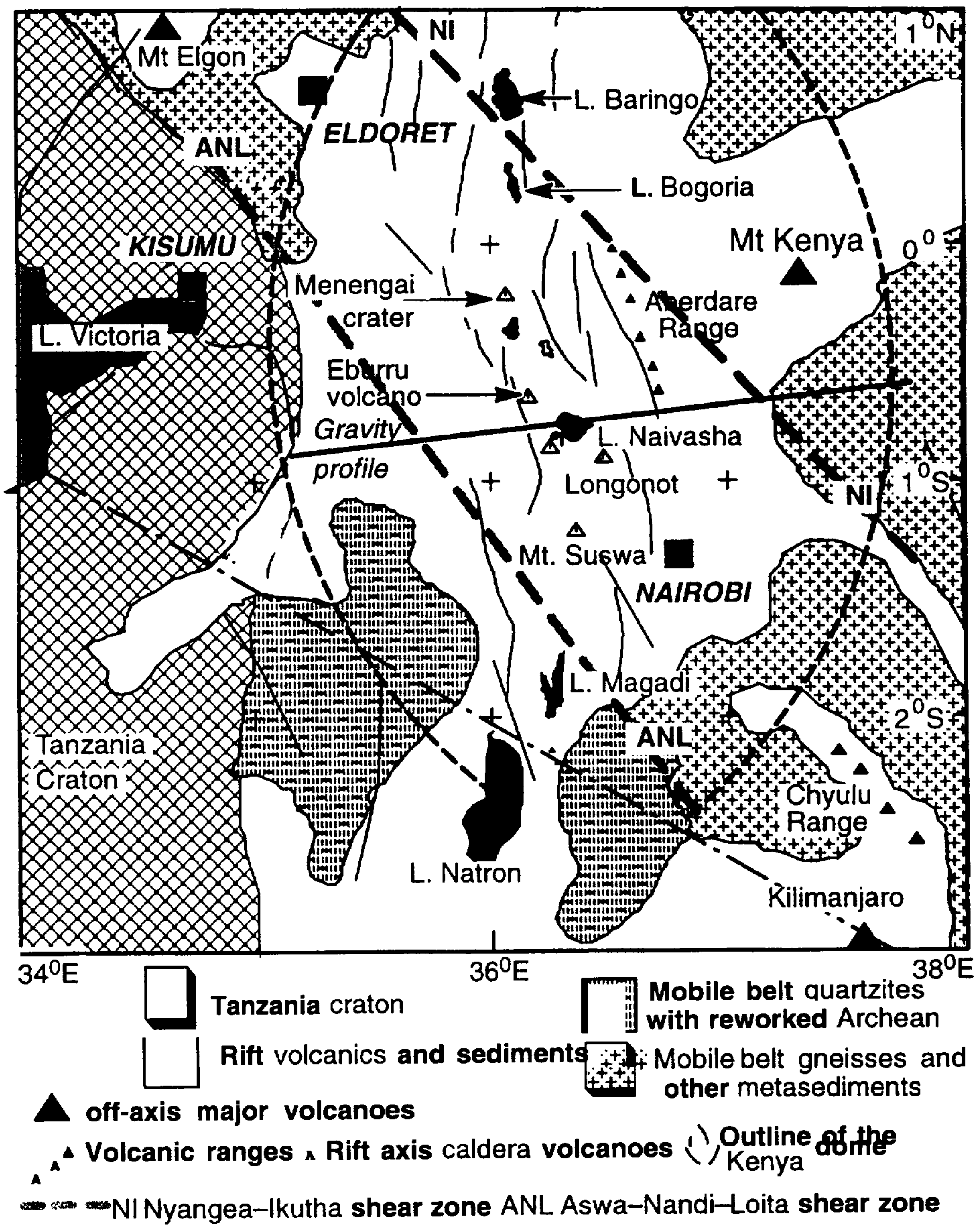


Figure 1: index map of southern Kenya rift showing volcanoes, faults, and gravity profile shown in Figure 2.

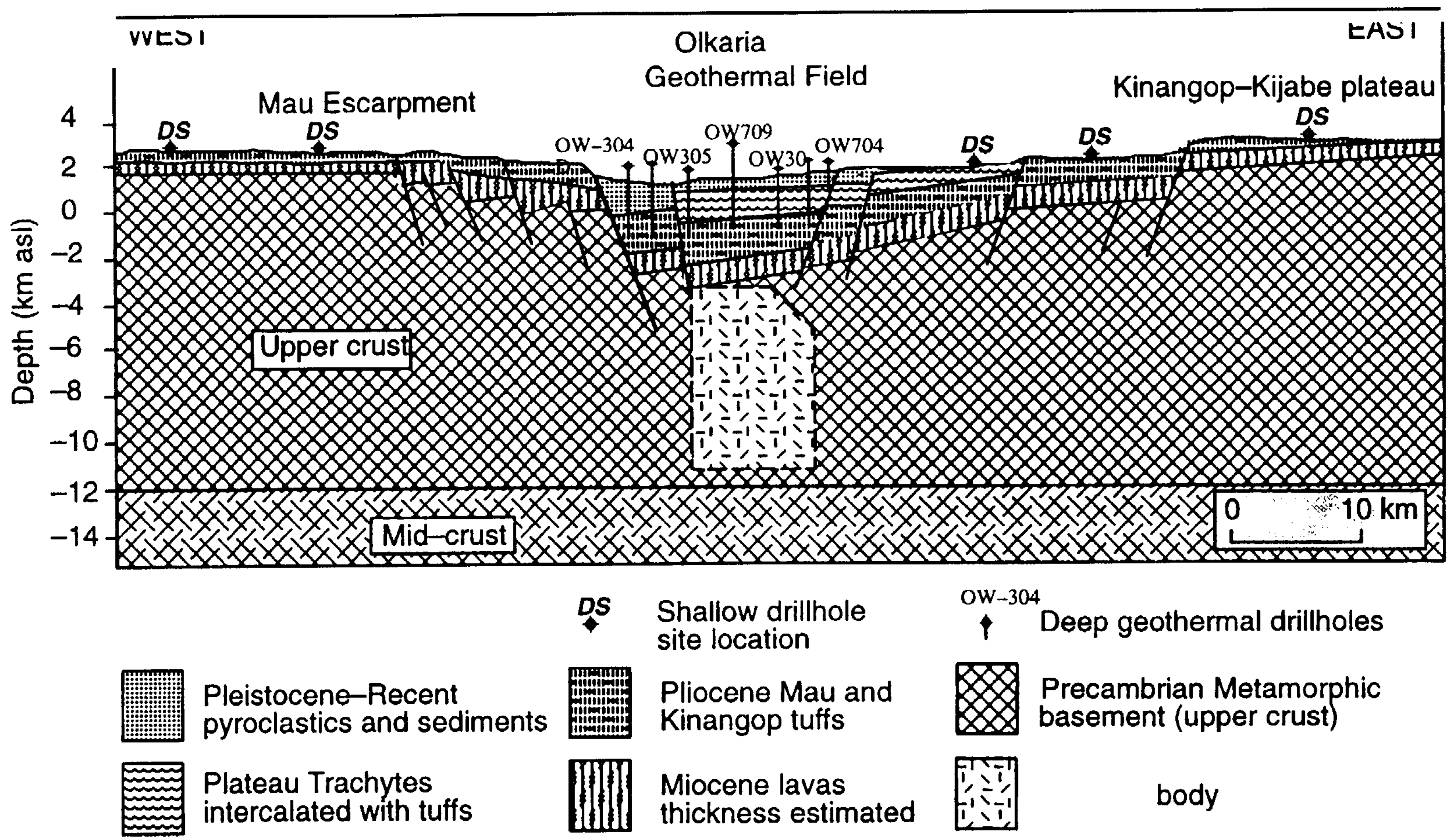


Figure 2: Interpreted cross–section along the gravity profile shown in Figure 1

URC97073

-72/64

Optimal Approximation of Quadratic Interval Functions

Misha Koshelev[1] and Patrick Taillibert[2]

[1]1003 Robinson
El Paso, TX 79902, USA
email mkosh@cs.utep.edu

[2] Dassault Electronique
55 Quai Marcel Dassault
92214 Saint Cloud, France
email Patrick. Taillibert@dassault-elec.f r

Abstract

Measurements are never absolutely accurate; as a result, after each measurement, we do not get the *exact* value of the measured quantity; at best, we get an in *terval* of its possible values.

For dynamically changing quantities x, the additional problem is that we cannot measure them continuously; we can only measure them at certain discrete moments of time t_1, t_2, ... If we know that the value $x(t_j)$ at a moment t_j of the last measurement was in the interval $[x^-(t_j), x^+(t_j)]$, and if we know the upper bound D on the rate with which x changes, then, for any given moment of time t, we can conclude that $x(t)$ belongs to the interval $[x^-(t_j) - D\,.\,(t - t_j), x^+(t_j) + D\,.\,(t - t_j)]$. This interval changes *linearly* with time, an is, therefore, called a *linear interval* function.

When we process these intervals, we get an expression that is quadratic and higher order w.r.t. time t. Such "quadratic" intervals are difficult to process and therefore, it is necessary to approximate them by linear ones.

In this paper, we describe an algorithm that gives the optimal approximation of quadratic interval functions by linear ones.

1 Introduction: intervals, linear and quadratic interval functions, and why it is necessary to approximate

The need for indirect measurements. In many real-life problems, we are interested in the values of a quantity y that is difficult or impossible to measure directly. For example:

- in *astrophysics, we* cannot directly measure the mass or the temperature of the star;
- in *geophysics, we* cannot directly measure the amount of oil in a given area, etc.

In all these situations, we measure some related quantities xl, ... x_n that can be measured directly, and then use the known relationship y = $f(x_1, \ldots, x_n)$ between x_i and y and the results $\tilde{x}_i$ of measuring x_i to estimate y as $\tilde{y} = f(\tilde{x}_1, \ldots, \tilde{x}_n)$. For example:

- to measure the temperature of the star, we measure its brightness $x_1, \ldots, x_n$ on different wavelengths, and then try to fit the resulting spectrum into a black-body radiation spectrum corresponding to the temperature y;
- to estimate the amount of oil y in a given area, we measure the geophysical characteristics of different parts of it, and use known equations to estimate y.

In contrast to a direct measurement of x_i, this two-stage process of direct measurements followed by data processing is called *indirect measurement.*

In space sciences, and especially in their applications to environmental and earth sciences, indirect measurements are the main method of measurement.

The main problem: estimating accuracy of the result of data processing. After the data processing is over, the main problem is: how accurate is the result $\tilde{y}$ of data processing?

For example, suppose that we estimate the amount of oil in a given area, and we got $\tilde{y} = 100$ mln ton. Our further actions will depend on the accuracy of this result:

- if the accuracy is high, e.g., if the actual amount of oil is 100 ± 1, then we will probably start drilling;
- however, if the accuracy is low, and the actual value if, say, 100 ± 100, then it may make more sense to undertake further, more accurate, measurements.

The inaccuracy in y comes from two sources:

- First, the model $y = f(x_1, \ldots, x_n)$ that is used in data processing is often inaccurate. This inaccuracy is the easiest to take into consideration: if we know the upper bound Δ_m on the error $y - f(x_1, \ldots, x_n)$ of this model, and if we know the values of x_i, then we can conclude that the actual value of y belongs to the interval $[\tilde{y} - \Delta_m, \tilde{y} + \Delta_m]$, where $\tilde{y} = f(x_1, \ldots, x_n)$.
- Second, the measurement results $\tilde{x}_i$ are also inaccurate and therefore, may differ from the actual values of the measured quantities x_i. This error is much more difficult to estimate, and this is what we will be doing in the present paper.

Intervals. Measurements are never 100% precise. Thus, if, as the result of measuring a certain quantity x;, we get a measurement result $\tilde{x}_i$, it does not necessarily mean that the actual value x_i of this quantity is exactly equal to $\tilde{x}_i$.

The manufacturer of the measuring instrument usually supplies it with the upper bound Δ_i for the measurement error $\Delta x_i = \tilde{x}_i - x_i$; in order words, the manufacturer guarantees that $|\Delta x_i| \leq \Delta_i$.

In this case, if we have measured a quantity x_i and the measurement result is $\tilde{x}_i$, then the only information that we have about the actual value is that this actual value cannot differ from $\tilde{x}_i$ by more than Δ_i, i.e., that this actual value must be within the *interval* $[\tilde{x}_i - \Delta_i, \tilde{x}_i$ + As].

If we know these intervals, then the set Y of possible values of $y = f(x_1, \ldots, x_n)$ can be described as follows:

$$Y = \{ f(x_1, \ldots, x_n) \mid x_1 \in [x_1^-, x_1^+], \ldots, x_n \in [x_n^-, x_n^+] \}. \tag{1}$$

Comments.

- Sometimes, in addition to the upper bound for the error, we know the *probabilities* of different error values. However, in many real-life cases, we do not know these probabilities, and the upper bound Δ_i is the only information about the measurement error Δx_i that we have.
- Computations that take this interval uncertainty into consideration are called ***interval computations*** *(see, e.g.,* [1, 2]).

Linear interval functions. In some cases, the values x_i do not change with time. In these cases, the only inaccuracy is caused by the inaccuracy of the direct measurement.

However, in many real-life situations, the measured values change with time. For such dynamically changing quantities xi, there is an additional source of uncertainty: namely, we cannot measure a quantity continuously; we can only measure it at certain discrete moments of time $t_1, t_2, \ldots$ Hence:

- If we are interested in the value of $x_i(t_j)$ at one of the moments of time t_j in which we have actually measured x_i, then the inaccuracy of the direct measurement is the only source of the measurement error, and we get the interval $[x_i^-(t_j), x_i^+(t_j)]$ of possible values of $x_i(t_j)$.
- However, if we are interested in the value xi(t) of the quantity x_i at a moment of time t in which no direct measurement of xi was performed, then we must use one of the measured values, and get an additional error component caused by the change of x_i between the moment t_j of the last measurement and the desired moment t. There are two methods to take this additional error into consideration:

  If we know the upper bound D_i on the rate with which z changes, then, for any given moment of time t, *we* can conclude that $x_i(t_j) - D_i \cdot (t - t_j) \leq x_i(t) \leq x_i(t_j) + D_i \cdot (t - t_j)$. Hence, from the fact that $x_i^-(t_j) \leq x_i(t_j) \leq x^+(t_j)$, *we* conclude that the value $x_i(t)$ belongs to the interval $[x_i^-(t), x_i^+(t)]$, where z;(t)= $x_i^-(t_j) - D_i \cdot (t - t_j)$ and $x_i^+(t) = x_i^+(t_j) + D_i \cdot (t - t_j)$.

- In some cases, simultaneously with measuring the values $x_i(t_j)$, we can directly measure the rate $\dot{x}_i(t_j)$ with which these values change; e.g.,
  * simultaneously with measuring coordinates $x_i(t)$ of a robot, we can measure its velocity $\dot{x}_i(t)$;
  * simultaneously with measuring the velocity x_i @), we can measure the acceleration $\dot{x}_i(t)$; etc.

  In these case, if we know the interval of possible values $[\dot{x}_i^-(t_j), \dot{x}_i^+(t_j)]$ of the rate, and we know that the upper bound S_i on the second time derivative, we can conclude that:
  * for all $t \in [t_j, t_{j+1}]$, the values $\dot{x}_i(t_j)$ belong to the intervals $[v^-, v^+]$, where $v^- = \dot{x}_i^-(t_j) - S_i(t_{j+1} - t_j)$ and $v^+ = \dot{x}_i^+(t_j) + S_i \cdot (t_{j+1} - t_j)$, and therefore, that
  * for each of these moments of time, the actual values of xi(t) belong to the interval $[x^-(t), x^+(t)]$, where Z-(t)= $x_i^-(t_j) - v^- \cdot (t - t_j)$ and Z+(t)= $x_i^+(t_j) + v^+ \cdot (t - t_j)$.

In general, the only information that we have about the actual value of $x(t)$ is that $x(t) \in$ x(t) = [x-(t), $x^+(t)$]. In other words, instead of a single *interval*, we have an ***interval function*** that assigns an interval $\mathbf{x}(t)$ to each moment of time t.

In our cases, the endpoints $x^-(t)$ and z'(t) of this interval are linear functions and therefore, it is natural to call this interval function *linear interval function.*

For linear data processing, linear interval functions remain linear. If the algorithm $f(x_1, \ldots, x_n)$ is a linear function, then, as a result of applying this algorithm to the interval linear functions, we get an estimate (1) that is either a linear or a piece-wise linear function of time t.

In [3, 4, 5], such linear interval functions were effectively used in electrical and electronic engineering analysis of circuits with linear elements.

For non-linear data processing, we need an approximation. In most real-life situations, however, data processing algorithms are non-linear. As a result, if we start with the intervals for xi(t) that are linear in time t, *we* end up with intervals for $y(t)$ whose dependence on t is much more complicated. For example:

- if we ***add*** or ***subtract*** two intervals that are linear in time, we still get the result that is linear in time; but
- if we ***multiply*** two intervals that are linear in time, then their endpoints also get multiplied, and, as a result, we get an interval function $[y^-(t), y^+(t)]$ in which the endpoints are ***quadratic*** functions of time (i.e., we get ***quadratic interval functions).***

If we multiply more, we get cubic, quartic, etc. functions. In principle, there is nothing wrong with this complexity, except for the fact that while a linear interval function requires only 4 numbers to store (2 coefficients of the lower endpoint and 2 coefficients of the upper endpoint), quadratic functions require 6 coefficients, cubic functions require 8, etc. The more complicated the function becomes, the more memory we need to store these coefficients, and the longer it takes to process these functions.

Thus, since we are often limited both in processing time and in memory (especially if the processing is done in an on-board computer), we must ***approximate*** the given complicated interval function by a simpler one.

In other words, if we have an interval function $\mathbf{y}(t) = [y^-(t), y^+(t)]$ that is known to contain the actual value y(t), we want to be able to find a simpler interval function $\mathbf{z}(t) = [z^-(t), z^+(t)]$ that for each t, contains the entire interval *y(t)* and is, thus, guaranteed to contain the actual value $y(t)$.

Of course, this approximation comes at a trade-of we simplify the expression, but we make the interval wider (and therefore, lose some information). Therefore, the narrower the approximating interval function, the better.

What we are planning to do. The simplest approximation problem of this type is the problem of approximating a quadratic interval function by linear ones. In this paper, we will present an optimal solution to this problem.

2 When is the approximation optimal? Mathematical formulation of the problem

Definition 1. *Let an interval $[t^-, t^+]$ be fixed.*

- *By an interval junction, we* mean a *mapping* **x** that *puts* into correspondence to each *number* $t \in [t^-, t^+]$ *an interval* $x(t) = [x^-(t), x^+(t)]$.
- *If both functions* $x^-(t)$ *and* $x^+(t)$ are *linear* functions oft (i.e., if $x^\pm(t) = x_0^\pm + x_1^\pm \cdot t$), *we say that the* interval *function is linear.*
- *If both* functions $x^-(t)$ and $x^+(t)$ are *quadratic functions oft (i.e., if* $x^\pm(t) = x_0^\pm + x_1^\pm t + x_2^\pm \cdot t^2$), **we** say that *the interval function is quadratic.*
- *We* say that a *linear* interval function $\mathbf{z}(t) = [z^-(t), z^+(t)]$ *approximates a* quadratic interval function $y(t) = [y^-(t), y^+(t)]$ *if* $y(t) \subseteq z(t)$ *for all t.*

Comment. The narrower the intervals, the better. So, our goal is to minimize the worse-case width of the approximating interval, i.e., the value $W(z) = \max_t(z^+(t) - z^-(t))$.

We will see that in some cases, for a given quadratic interval function $\mathbf{y}(t)$, there are several approximating linear interval functions z with the same value of W(z). If two different approximating functions have the same worst-case widths, then it is reasonable to choose the one for which the best-case width $w(z) = \min_t(z^+(t) - z^-(t))$ is the smallest. Thus, we arrive at the following definition:

Definition 2.

- For every *interval function z(t):*
  - *by its worst-case width, we mean a* value
$$w(z) = \max_{t\in[t^-,t^+]}(z^+(t) - z^-(t)).$$
  - *by its best-case width, we* mean a *value*
$$w(z) = \min_{t\in[t^-,t^+]}(z^+(t) - z^-(t)).$$
- *Let* a quadratic interval function y be *fixed. We say that* a *linear* function z *is an optimal* approximation *of y if the following* conditions are satisfied:
  - first, z is an approximation of y.
  - second, among *all* linear approximations *to y, the* function z *has the* smallest value *of the worst-case width W(z);*
  - *third,* if *there exist several linear approximations* **u** to y, *with the same smallest* value *of the* worst-case *width* **W'(u),** *the function z has the* smallest *value of the best-case width w(z).*

What we are planning to do. In this paper, we will describe the optimal linear approximation to an arbitrary quadratic interval function.

The resulting formulas will depend on the signs of the coefficients (y_2^- and y_2^+) at the quadratic term t^2. In each of the resulting four cases, we will describe the geometric idea that leads to the optimal approximation, and give the explicit formulas for this approximation.

3 Case 1: $y_2^- \leq 0$, $y_2^+ \geq 0$

Derivation of the solution. In this case, both functions $y^-(x)$ and $y^+(z)$ are "pointing inward".

If $\mathbf{z}(t)$ *is* a linear interval approximation to the given quadratic interval function, then, from $\mathbf{y}(t) \subseteq z(t)$, it follows, in particular, that $z^-(t) \leq y^-(t)$ for all $t \in [t^-, t^+]$; therefore, $z^-(t^-) \leq y^-(t^-)$ and $z^-(t^+) \leq y^-(t^+)$.

Since $y_2^- \leq 0$, the quadratic function $y^-(t)$ is concave and hence, from $z^-(t^-) \leq y^-(t^-)$ and $z^-(t^+) \leq y^-(t^+)$, *i.e.,* from the fact that the line $z^-(t)$ lies below the two endpoints of the graph of $y^-(t)$, it automatically follows that the entire graph of $y^-(t)$ is above the line $z^-(t)$. So, it is sufficient to guarantee that at the endpoints, the values $z^-(t^\pm)$ do not exceed the corresponding values of $y^-(t^\pm)$.

For fixed $z^+(t)$, the resulting intervals are the narrowest when the line z (t) is at the highest possible location. Thus, to minimize the widths of the intervals, we must move both points $z^-(t^\pm)$ up as much as possible. The highest possible location for $z^-(t^-)$ is y (t), and the highest possible location for $z(t^+)$ is $y^-(t^+)$. Thus, y-(t) is a a straight line going through $y^-(t^\pm)$, i.e., a secant.

Similarly, z^+ (t) is a secant of $y^+(t)$.

Solution. In the optimal approximation, the function $z^-(t)$ is the *secant* of y^- (t), i.e., a straight line whose endpoints are the endpoints of the quadratic function y (t) on this interval. Similarly, z'(t) is a secant of $y^+(t)$, i.e.,

$$z^-(t) = y^-(t^-) + \frac{y^-(t^+) - y^-(t^-)}{t^+ - t^-} \cdot (t - t^-) \tag{2}$$

$$z^-(t) = y^+(t^-) + \frac{y^+(t^+) - y^+(t^-)}{t^+ - t^-} \cdot (t - t^-). \tag{3}$$

4 Case 2: $y_2^- < 0$, $y_2^+ < 0$

Derivation of the solution. The arguments given for Case 1 show that in this case, the function $z^-(t)$ is still the secant. With z (t) fixed, it is sufficient to find the upper function *z+(t)* for which the resulting approximation is optimal.

Let us first guarantee that the worst-case width is indeed the smallest possible. If this width is equal to W_0, this means that $z^+(t) \le z(t) + W_0$, and therefore, that $y^+(t) \le z^+(t) \le z^-(t) + W_0$. Thus, to guarantee that W_0 takes the smallest possible values, we must choose W_0 as the smallest possible value for which y'(t) $\le z(t) + W_0$ for all $t \in [t^-, t^+]$. In other words, as W'., we take the maximum of the function $y^+(t)$ - *z-(t)* on the interval $[t^-, t^+]$.

The maximum of the concave quadratic function y'(t) - $z(t) = (y_0^+ - z_0^-) + (y_1^+ - z_1^-) \cdot t + y_2^- t^2$ is attained at the point where its derivative is equal to O, i.e., at the point

$$t_m = -(y_1^+ - z_1^-)/(2y_2^+). \tag{4}$$

If this maximum is attained at the internal point t_m of this interval, then at t_m, the line $z^-(t) + W_0$ is a tangent to y+(t), and therefore, there is no way to find a lower line without increasing the worst-case width $W(\mathbf{z})$. So, in this case, $z^+(t)$ = *z-(t)+* W_0.

If the maximum is attained in one of the endpoints, e.g., at t^+, then, we can, keeping the straight line z^+ (t) at the point (t+, y^+ (t+)), lower its other end and still get the same worst-case width. The lowest value of the best-case width is attained when we lower the other end to the lowest possible position in which it is still above $y^+(t)$, i.e., to the position of a *tangent* to $y^+(t)$.

Solution. For the optimal approximation, the lower line $z^-(t)$ is a secant (2). To determine the upper line $z^+(t)$, we apply the formula (4) to compute the value t_m. Then:

- If $t_m \in$ *[t–, t+], we* take $z^+(t) = z^-(t) + (y^+(t_m) - z^-(t_m))$.

o If $t_m > t^+$, then z^+ is the tangent to y^+ at t^+:

$$z^+(t) = y^+(t^+) + (y_1^+ + 2y_2^+ \cdot t^+)(t - t^+). \tag{5}$$

o If $t_m <$ t, then z^+ is the tangent to y^+ at t:

$$z^+(t) = y^+(t^-) + (y_1^+ + 2y_2^+ \cdot t^-)(t - t^-). \tag{6}$$

5 Case 3: $y_2^- > 0$, $y_2^+ > 0$

This case is similar to case 2, so we can immediately give a solution:

Solution. For the optimal approximation, the upper line $z^+(t)$ *is* a secant (3). To determine the lower line *z–(t), we* compute the value $t_m = -(y_1^- - z_1^+)/(2y_2^-)$. Then:

o If $t_m \in [t^-, t^+]$, *we* take *z-(t)* = $z^+(t) - (z^+(t_m) - y^-(t_m))$.

- If $t_m > t^+$, then z is the tangent to y at t^+:

$$z\text{-}(t) = y^-(t^+) + (y_1^- + 2y_2^- t^+)(t - t^+). \tag{7}$$

- If $t_m < t^-$, then z^- is the tangent to y at t^-:

$$z^-(t) = y^-(t^-) + (y_1^- + 2y_2^- t^-)(t - t\text{-}). \tag{8}$$

6 Case 4: $y_2^- < 0,\ y_2^+ < 0$

Derivation of the solution. In this case, both functions $y^-(x)$ and y'(x) are "pointing outward".

As a first step of constructing the optimal linear approximation z(t), let us first make sure that we have the smallest possible value of the worst-case width. For every $t \in [t^-, t^+]$, from $\mathbf{y}(t) \subseteq z(t)$, *we* can conclude that the width of z(t) is at least as large as the width of the interval $\mathbf{y}(t)$. Thus, the worst-case width W(z) of the approximating linear function z(t) cannot be smaller than the worst-case width W(y) of the original (quadratic) interval function y. Since we are minimizing W'(z), it is therefore desirable to choose z in such a way that its worst-case width is ***exactly equal*** to ***W(y).***

The worst-case width *W(y) is a* maximum of the quadratic width function $y^+(t) - y(t)$ on the interval $[t, t^+]$. By differentiating this difference, one can easily get an explicit expression for this maximum point t_M (see below).

If this maximum point t_M is inside the interval $[t, t^+]$, then at this point t_M, both approximating lines Z(t) and $z^+(t)$ must be tangent to the corresponding functions $y^-(t)$ and y+(t), because otherwise, in at least one of the directions, the width will increase.

If this maximum t_M is attained at one the endpoints, e.g., for t', then we must have $z^-(t^+) = y(t^+)$ and $z^+(t') = y^+(t')$. In this case, to guarantee the smallest possible ***best-case*** width, we must place z'(t) as *low* as possible (i.e., along the tangent to y^+), and $z^-(t)$ as high as possible, i.e., similarly, along the tangent to $y(t)$.

As a result, we arrive at the following formulas:

Solution. Compute $t_M = -(y_1^+ - y_1^-)/[2(y_2^+ - y_2^-)]$. Then:

- If $t_M \in [t^-, t^+]$, then z^- is the tangent to y^- at t_M, and z^+ is the tangent to y^+ at t_M: $z\text{-}(t) = y^-(t_M) + (y_1^- + 2y_2^- \cdot t_M)(t - t_M)$; $z^+(t) = y^+(t_M) + (y_1^+ + 2y_2^+ \cdot t_M)(t - t_M)$.
- If $t_M > t^+$, then z is the tangent (7) to y at t_+, and z' is the tangent (5) to y^+ at t_+;
- If $t_M < t^-$, then z^- is the tangent (8) to y at t_- and z^+ is the tangent (6) to y^+ at t_-.

Acknowledgments. This work was partially supported by the NASA Pan American Center for Environmental and Earth Studies (PACES). The authors are thankful to Ann Gates, Vladik Kreinovich, Luc Longpré, and Scott Starks for their help and encouragement.

References

[1] R. B. Kearfott, ***Rigorous global search: continuous problems,*** Kluwer, Dordrecht, 1996.

[2] R. B. Kearfott and V. Kreinovich (eds.), ***Applications of Interval Computations,*** Kluwer, Dordrecht, 1996.

[3] O. Lhomme, "Consistency Techniques for Numeric CSPs," In: ***Proceedings of IJCAI'93, the Thirteenth International Joint Conference on Artificial Intelligence, 1993.***

[4] O. Lhomme, A. Gottlieb, M. Rueher, and P. Taillibert, "Boosting the Interval Narrowing Algorithm," In: ***Proceedings of JICSLP '96, Joint International Conference and Symposium on Logic Programming,*** Bonn, ***Germany, September 2-6, 1996.***

[5] E. Loiez and P. Taillibert, "Analog Systems Diagnosis: Modeling with Temporal Bands," In: ***Proceedings of CESA '96, IEEE International Conference on Computations in Engineering and Simulations Applications, Line, France, July 9-12, 1996.***

URC97074

7,4 / 61

An Arbitrary First Order Theory Can Be Represented by a Logic Program: a Theorem

Olga Kosheleva[1,2]

[1] Department of Electrical and Computer Engineering
and [2]Knowledge Representation Laboratory
University of Texas at El Paso
El Paso, TX 79968
email olga@ece.utep.edu

Abstract

How can we represent knowledge inside a computer?

For formalized knowledge, classical logic seems to be the most adequate tool. Classical logic is behind all formalisms of classical mathematics, and behind many formalisms used in Artificial Intelligence.

There is only one serious problem with classical logic: due to the famous Gödel's theorem, classical logic is algorithmically undecidable; as a result, when the knowledge is represented in the form of logical statements, it is very difficult to check whether, based on this statement, a given query is true or not.

To make knowledge representations more algorithmic, a special field of logic programming was invented. An important portion of logic programming is algorithmically decidable. To cover knowledge that cannot be represented in this portion, several extensions of the decidable fragments have been proposed. In the spirit of logic programming, these extensions are usually introduced in such a way that even if a general algorithm is not available, good heuristic methods exist.

It is important to check whether the already proposed extensions are sufficient, or further extensions is necessary. In the present paper, we show that one particular extension, namely, logic programming with classical negation, introduced by M. Gelfond and V. Lifschitz, can represent (in some reasonable sense) an arbitrary first order logical theory.

1 Introduction

Intelligent **data processing is extremely important in space applications.** One of the main problems with space-related data processing is that the amount of data grows so fast that, by some estimates, only about 10% of the data is being processed.

We humans also get lots of information, but our brain is accustomed to filtering out the irrelevant information and processing only the relevant one. To use this experience, we need to use intelligent data processing techniques.

For that, we must be able to *represent* our knowledge in the computer in such a way that we will be able to use this knowledge for processing data.

Classical logic is the natural way of representing human knowledge, but classical logic is non-algorithmic. How can we represent knowledge inside a computer?

For *formalized* knowledge, the most adequate tool seems to be *classical logic* (see, e.g., [6, 2, 1]). Classical logic is behind all formalisms of classical mathematics, and behind many formalisms used in Artificial Intelligence.

There is only one serious problem with classical logic: due to the famous Gödel's theorem, classical logic is algorithmically undecidable; as a result, when the knowledge is represented in the form of logical statements, it is very difficult to check whether, based on this statement, a given query is true or not.

Logic programming: an attempt to make logic algorithmic. To make knowledge representations more algorithmic, a special field of *logic* programming was invented.

Extensions of traditional logic programming. An important portion of logic programming is algorithmically decidable.

To cover knowledge that cannot be represented in this portion, several extensions of the decidable fragments have been proposed.

In the spirit of logic programming, these extensions are usually introduced in such a way that even if n general algorithm is not available, good heuristic methods exist.

An important problem: are the existing extensions sufficient? It is important to check whether the already proposed extensions are sufficient, or further extensions is necessary.

What we are planning to do. In the present paper, we show that one particular extension, namely, logic programming with classical negation, introduced by M. Gel fond and V. Lifschitz [3, 4], can represent (in some reasonable sense) an arbitrary first order logical theory.

Moreover, we will capitalize on the fact that logic programming can describe *transitive closure* that cannot be represented in traditional first order logic, and show that this logic programming formalism can describe extensions of first order theories obtained by adding this notion of a transitive closure.

The preliminary results of this paper first appeared as a draft [5].

The structure of this paper. To make this result more accessible to general readers, we will briefly recall the main definitions of classical logic and of logic programming with classical negation.

2 Basic definitions

2.1 Classical (first-order) logic: a reminder

Definition 1. Suppose that we are given three sets C, $\mathcal{V}$, and $\mathcal{P}$ with $|\mathcal{C}| \leq$ No, $|\mathcal{V}| \leq$ No, and $|\mathcal{P}| \leq$ No, and a function *ar* from $\mathcal{P}$ to the *set N of* non-negative integers.

- *Elements of the set C will be called constants and denoted by* $c_1, \ldots, c_n, \ldots$
- Elements *of* the set V are called *variables and* denoted by $x_1, \ldots, x_n, \ldots$
- *Elements of $\mathcal{P}$ will be called predicate symbols and denoted by* $P_1, \ldots, P_n, \ldots$
- The value $ar(P_i)$ will be *called the arity of* a *predicate* P_i:
  - a predicate of *arity* 1 is *called unary;*
  - a predicate of *arity 2 is* called binary;
  - a predicate of *arity* 3 is *called ternary;*
  - *etc.*
- *By* an *atom, we* mean an expression *of* the type $P(x, \ldots, y)$, *where* $P \in \mathcal{P}$, each of the *symbols* $x, \ldots, y$ is either a *constant* or a variable, and the number of these symbols x, ... y coincides with the *arity* of the predicate *symbol P.*
- If *all* the *symbols* $x, \ldots, y$ *in* the definition of an atom are constants, then *this* atom is *called* a ground *atom.*
- *By* a *first order formula we mean a* closed formula A that is formed from atoms by using logical connective ($\vee$, &, $\neg$, $\rightarrow$, $\equiv$) and quantifiers $\forall x_i$ and $\exists x_i$.
- *By* a *first* order *theory* T *we will* mean a finite set of first order *formulas* $\{A_1, \ldots, A_t\}$.

Definition 2. *For an arbitrary first order theory, we can define a model as a set U (called a Universe), and relations* $\tilde{P}_i$ *on* this set U (for *all* P_i that occur in T) that satisfy *all* the formulas A_j *from the theory* T.

We say that a formula F follows from T, and denote it $T \models F$, if F is true in all *models* of T.

2.2 Adding transitive closure (TC) to the first order logic

Definition 3.

- By a *TC-formula* we mean either a (closed) first order formula, or an expression of the type $TC(P_i, P_j)$, where P_i and P_j are binary predicates.
- By a *TC-theory* we mean a finite set of TC-formulas $\{A_1, \ldots, A_t\}$.
- For a TC-theory T, by its *first order part* $\tilde{T}$, we mean the set of all first order formulas from T.
- By a *model* of a TC-theory T we mean such a model of $\tilde{T}$, that if $TC(P_i, P_j) \in T$, then $\tilde{P}_i$ is a transitive closure of $\tilde{P}_j$.
- If a formula F is true in all models of a TC-theory T, then we say that F *follows from* T, and denote it by $T \models F$.

2.3 Facts and queries

Motivations. Each theory represents a general description of the objects that we are interested in. E.g., it may describe a linear ordering. To be more specific, we must add some knowledge about our specific object. This knowledge is usually presented in the form of facts, i.e., atomic statements.

After we add this knowledge, we may ask whether some basic statement is true for the resulting theory or not. So, we arrive at the following definition:

Definition 4.

- By a *fact* we mean a ground atom or its negation. Facts will be denoted by $F_1, \ldots, F_n, \ldots$
- By a *query* we also mean a ground atom or its negation. Queries will be denoted by Q.

Comment. In logic, what, we call a fact, or a query, is usually called a *literal*

2.4 Definitions of generalized logic programs: a reminder

In the present paper, we consider logic program with classical negation in the sense of [3, 4].

We want to formulate logic programs that are equivalent to first order theories. It turns out that for that purpose, we must use additional (auxiliary) constants, predicates and functional symbols. So, we arrive at the following definitions:

Definition 5. Suppose that in addition to the sets C, V, and P, we have denumerable sets $\mathcal{R}$, $\mathcal{B}$, and $\mathcal{F}$, and a function $arity: \mathcal{F} \to N$ such that for every $n \in N$, there are infinitely many $f \in \mathcal{F}$ with $arity(f) = n$,

- Elements of the set $\mathcal{B}$ will be called *auxiliary constants* and denoted by $b_1, \ldots, b_n, \ldots$
- Elements of the set $\mathcal{R}$ will be called *auxiliary predicates* and denoted by $R_1, \ldots, R_n, \ldots$
- Elements of $\mathcal{F}$ will be called *auxiliary functional symbols* and denoted by $g, \ldots, g_n, \ldots$ For each $f \in \mathcal{F}$, the value $arity(f)$ is called an *arity* of f.
- A *term* is defined in the usual manner, starting from constants, auxiliary constants and variables, and applying function symbols of appropriate arity.
- By a *generalized atom* we mean an expression of the type $P(t_1, \ldots, t_n)$, where $P \in \mathcal{P} \cup \mathcal{R}$ is a predicate or auxiliary predicate of arity n, and t_i are terms.
- A *generalized literal* is a generalized atom p or the expression of the type $\neg p$, where p is a generalized atom; an expression $\neg p$ is called *classical negation*.
- A *rule* is an expression of the type $A \leftarrow B_1, \ldots, B_m$, where A is a generalized literal, $m \geq 0$, and each of B_i is either a generalized literal, or an expression of the type not p for some generalized literal p,
- Rules with $m = 0$ are called *facts*. A fact $A \leftarrow$ can also be written as A.
- A finite set of rules is called a *generalized logic program*, or a *logic program with classical negation*. Such programs will be denoted by $\mathcal{P}$, $\mathcal{P}_i$, etc.

- We say *that a query Q is true for a* program 'P (and *denote* it by $P \mid\sim Q$) if Q belongs to any *consistent* answer set *of P (in the sense of [3, 4])*.

Comments

- Please note that in the formulation of the query, we only allow the symbols from the *original* theory, auxiliary symbols are not allowed.
- Since in this paper, we will only use logic programs with classical negation, we will call them, without confusion, simply *logic programs.*

3 Main result

THEOREM. *There exists an algorithm that transform every TC-theory T into a logic program $\mathcal{P}_T$ with classical negation so that for an arbitrary finite set of facts $\{F_1, \ldots, F_n\}$, and for an arbitrary query Q, Q is true in $T + \{F_1, \ldots, F_n\}$ if and only if Q it true in $\mathcal{P}_T + \{F_1 \leftarrow, \ldots, F_n \leftarrow\}$.*

Comment. We would like to emphasize once again that we allow the use of auxiliary predicates, constants and function symbols while describing the rules of the logic program, but not in queries or facts. So, in this Theorem, we still apply Definitions 4 to describe facts and queries. According to these definitions, facts and queries are ground atoms (or negations of ground atoms) that are formed only from the original predicate symbols P_i and original constants c,.

4 Description of the algorithm and the main idea of the proof

Let us describe the algorithm that transform a theory into *a* logical program.

4.1 Case of first order theories

At first, we will consider the case when the TC-theory does not contain any statements about the transitive closure, i.e., when it is actually the first order theory. We will illustrate this case on the example of the following theory that describes dense order:

$$\forall x, y, z\,(x < y \;\&\; y < z \rightarrow x < z);$$

$$\forall x, y\,(x < y \rightarrow \neg y < x);$$

$$\forall x\,(\neg x < x);$$

$$\forall x, y\,\exists z\,(x < z \;\&\; z < y).$$

Step 1: general description. First we make a *skolemization* of the axioms of the given first-order theory (for definitions, see, e.g., [6, 2, 1]).

Step 1: example. In our example, skolemization leads to the following axioms (universal quantifiers are, for simplicity, omitted):

$$x < y \;\&\; y < z \rightarrow x < z;$$

$$x < y \rightarrow \neg y < x;$$

$$\neg x < x:$$

$$x < f(x, y) \;\&\; f(x, y) < y.$$

Step 2: general description, Every axiom is represented in conjunctive normal form [6, 2, 1]), and each of the resulting conjunctions is written separately.

Step 2: example. In our example, we will get the following set of disjunctions:

$$(\neg x < y) \vee (\neg y < z) \vee (x < z);$$

$$(\neg x < y) \vee (\neg y < x);$$

$$(\neg x < z);$$

$$(x < f(x, y));$$

$$(f(x, y) < y).$$

Step 3: general description. On this step, we translate every disjunction al V . . . Van into the following n rules:

$$an \leftarrow \neg a_1, \neg a_2, \ldots, a_{n-1}.$$

$$a_{n-1} \leftarrow \neg a_1, \neg a_2, \ldots, a_{n-2}, a_n.$$

$$\ldots$$

$$a_1 \leftarrow \neg a_2, \ldots, a_n.$$

Step 3: example. In our example, we will get the following program (in classical logic, it is a usual practice to have a predicate symbol like < in between the arguments, but in logic programming, the predicate symbol is usually in front; *to* follow this tradition, we will use a notation $L(x, y)$ instead of $x < y$):

$$L(x, z) \leftarrow L(x, y), L(y, z).$$

$$\neg L(y, z) \leftarrow L(x, y), \neg L(x, z).$$

$$\neg L(x, y) \leftarrow L(y, z), \neg L(x, z).$$

$$\neg L(x, y) \leftarrow L(y, x).$$

$$\neg L(y, x) \leftarrow L(x, y).$$

$$\neg L(x, x).$$

$$L(x, f(x, y)).$$

$$L(f(x, y), y).$$

Step 4: general description. Finally, to obtain a program $\mathcal{P}_T$ that is "equivalent" to the original theory T (in the sense of Theorem 1) we add, for each of the predicates $P(x, ..., y)$ from the resulting program, two statements called ***Closed World Assumption*** (CWA):

$$P(x, \ldots, y) \leftarrow \text{not} \neg P(x, \ldots, y).$$

$$\neg P(x, \ldots, y) \leftarrow \text{not } P(x, \ldots, y).$$

Step 4: example. In particular, in our example, we add the following two statements:

$$L(x, y) \leftarrow \text{not } \neg L(x, y).$$

$$\neg L(x, y) \leftarrow \text{not } L(x, y).$$

Idea of the proof. We need to prove that for any finite set of facts $\{F_1, \ldots F_n\}$, and for an arbitrary query Q, Q is true in $T + \{F_1, \ldots F_n\}$ if **and** only if Q it true in $\mathcal{P}_T + \{F_1 \leftarrow, \ldots, F_n \leftarrow\}$.

For classical logic, Q is true in $T + \{F_1, \ldots F_n\}$ iff the theory $T' = T + \{F_1, \ldots F_n\} + \neg Q$ *is* **inconsistent. The inconsistency of the theory is equivalent to the inconsistency of its** skolemization, **so, it is sufficient** to check whether the skolemized version $S(T')$ is inconsistent, i.e., whether Q is deducible from the theory $T'' = S(T) + \{F_1, \ldots F_n\}$, i.e., whether Q is true in all models of T''. It is sufficient to consider Herbrand models of T''.

It is easy to show that every Herbrand model of T'' is a consistent answer set of the corresponding logic program (minimality follows from the present of the two close world assumptions), and vice versa, every consistent answer set represents a Herbrand model of T''. This observation concludes the proof.

4.2 Theories with transitive closure

If a theory T contains statements about transitive closure, then we need to add the following additional step to our algorithm:

Step 5. If the original theory T contains the expression $TC(A, B)$ for some binary predicate symbols A and B, then we:

- add an auxiliary predicate α_{AB} to the set $\mathcal{R}$ of auxiliary predicates; and
- acid the following rules to the logic program obtained on Step 4:

$$\alpha_{AB}(x, y) \leftarrow B(x, y).$$

$$\alpha_{AB}(x, y) \leftarrow B(x, y), \alpha_{AB}(y, z).$$

$$A(x, y) \leftarrow \alpha_{AB}(x, y).$$

$$\neg A(x, y) \leftarrow \neg\alpha_{AB}(x, y).$$

$$\neg\alpha_{AB}(x, y) \leftarrow \text{not } \alpha_{AB}(x, y).$$

Comment. It is easy to show that the "standard" way of representing transitive closure in logic programming will not work. Indeed, traditionally, the fact that predicate anc (ancestor) is a transitive closure of the predicate par (parent) is expressed as follows:

$$anc(x, y) \leftarrow par(x, y).$$

$$anc(x, y) \leftarrow par(x, z), anc(z, y).$$

$$\neg anc(x, y) \leftarrow \text{not } anc(z, y).$$

However, if we add the facts

$$\neg par(a_i, a_j) \leftarrow$$

for all i, j, and

$$anc(a_1, a_2) \quad +$$

then we get P + $F \mid\!\sim anc(a_1, a_2)$, but in this model, the transitive closure par^* is empty, and is, therefore, different from ans.

4.3 General comment

The proof given above shows that the correspondence between theories and logic programs is even more straightforward that follows from our Theorem: Namely, if we add a new axiom A_{t+1} to the theory T, then a logic program that corresponds to the resulting theory $\tilde{T}$, can be obtained from T by adding rules that correspond to A_{t+1}.

Acknowledgments. This work was partly supported by the NASA Pan American Center for Environmental and Earth Studies (PACES). The author is thankful to Chitta Baral, Ann Gates, Michael Gelfond, Vladik Kreinovich, Luc Longpré, Arthur Ramer, and Scott Starks for their help and encouragement.

References

[1] J. Barwise (cd.). ***Handbook of Mathematical*** *Logic*. North-Holland, Amsterdam, 1977

[2] H. B. Enderton. ***A mathematical introduction to logic.*** Academic Press, N. Y., 1972.

[3] M. Gelfond and V. Lifschitz, "Logic programs with classical negation", In: D. Warren and P. Szeredi (eds.), Logic *Programming: Proceedings of the 7th International* Conference, 1990, pp. 579-597.

[4] M. Gelfond and V. Lifschitz "Classical negation in logic programs and disjunctive databases", *New Generation Computing,* 1991, Vol. 9, pp. 365 385.

[5] O. Kosheleva, ***Any theory expressible in first order logic extended by transitive closure can be represented by a logic*** program, Draft, October 1992.

[6] J. R. Schoenfield. ***Mathematical*** *logic*. Addison-Wesley, 1967.

URC97075

RADIATION CHARACTERISTICS OF THE 486-DX4 MICROPROCESSOR

Coy Kouba & Gwan Choi
Department of Electrical Engineering
Texas A&M University
{ckouba, gchoi}@eesun1.tamu.edu

75/65

November 03, 1996

I. INTRODUCTION

This work describes the development of an experimental radiation testing environment to characterize the ***Single Event Upset*** susceptibility of the 486-DX4 microprocessor. Single event upsets (SEUs) are of particular concern to highly-reliable computer systems operating in radiation environments, such as in space or in nuclear reactors. A single event upset occurs when ionizing radiation strikes a sensitive junction in an operational integrated circuit [1]. An excess amount of energy is deposited at the junction, causing a change in its logic state (i.e., a "bit flip"). This upset is generally a soft error which can be corrected by reprogramming the affected location. The impact to the system however, can lead to data corruption or program flow anomalies, depending on the location and nature of the upset. They may either go completely unnoticed if the upset location is never used, or they may lead to catastrophic results if a critical register or logic function is altered [3].

Therefore, SEU testing and analysis techniques are needed to effectively evaluate a device's potential for reliability and resiliency before it is utilized in the harsh environment of space. The goal of this research was to develop an integrated SEU testing environment using a cyclotron facility and an advanced SEU monitoring system. The Texas A&M Cyclotron Institute was used to provide the radiation environment, with heavy ions being used as the source. The 486-DX4 microprocessor was the first device tested because it is currently under consideration for use in a data management system aboard the International Space Station. This research encompasses both experimental and analytical techniques, and yields a characterization of the 486-DX4's behavior for various operating modes.

II. TEST PHILOSOPHY & METHODOLOGY

A. Objectives

While previous radiation testing has generally focused on the device level, our approach is to test the device in an manner that is consistent with the actual application environment in which it is to be used. For the 486-DX4 testing, this includes using a PC-based system board with all associated peripherals to host the processor, however only the processor will be exposed to the radiation beam. This will allow the CPU (or any other system component) to be tested in an integrated fashion, just as it would be configured for operation in its intended environment. This same setup can also be used for testing other devices in future tests, such as memory modules, microcontrollers, etc.

The major objectives of the 486-DX4 cyclotron test were to: (1) establish the radiation upset characteristics for the 486-DX4 microprocessor, (2) calculate the upset capture cross-section of each device using several ion beams, (3) determine the SEU and latchup thresholds, (4) identify the error modes observed in each processor, and (5) test different chip implementations of the 486-DX4 using two different vendors. Two devices from each vendor were tested, and are referred to in this paper as Vendor A and Vendor B.

B. Hardware and Software Setup

To carry out our objectives, a comprehensive hardware& software test system had to be constructed. The hardware consisted of a 486 mother board with a custom-designed extender socket. This socket flipped the chip over on its back and allowed the radiation beam to directly bombard the CPU through a hole in the bottom of the socket. The lid of the CPU chip had to be removed in order to allow sufficient energy to penetrate the silicon die. Test data was acquired by using a video monitor and an HP16500B logic analyzer with modified signal acquisition cables. Test stimulus and control was achieved using the motherboard, a video monitor, keyboard, two power supplies, and the logic analyzer.

A current-limiting power supply was used to power the test board, and would be the primary means of detecting latchup conditions. The current limit was set at 100 mA above the nominal current draw. Separate power supplies were used to power the disk drives and the +/-12 volts and -5 volts on the test board. An external hard disk drive was connected outside the test chamber, and stored all test programs as well as the test data. A thermocouple was attached to the device under test and monitored at the test control station. If the temperature of the device ever exceeded 85 degrees Celsius, it would be powered down and allowed to cool. Refer to Figure 1 for an illustration of the cyclotron hardware test setup.

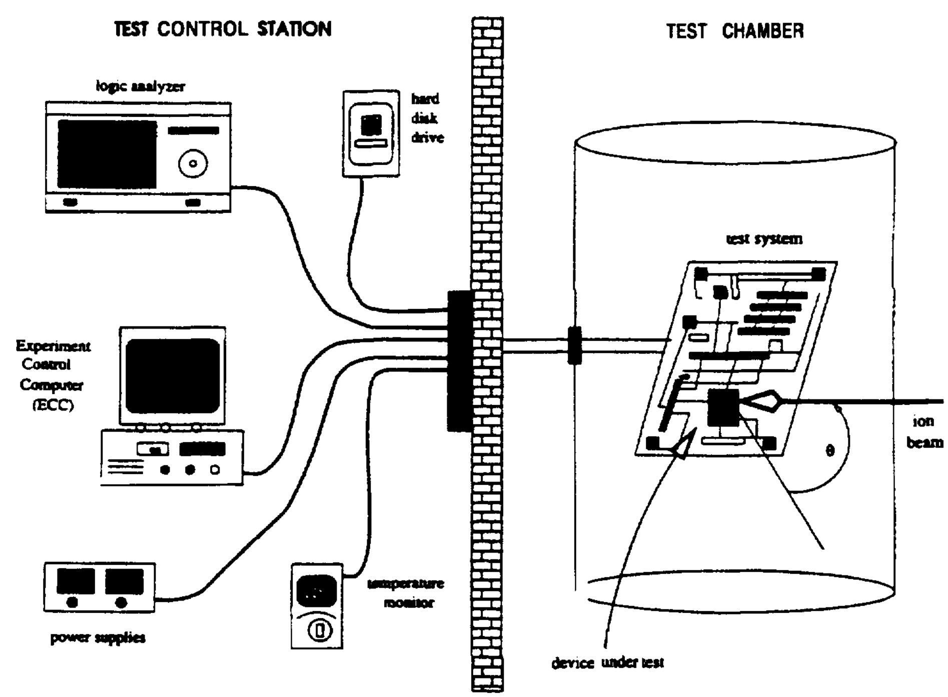


Figure 1 - Hardware setup for the radiation test

Six software programs were written to test the 486-DX4. Four programs were used to test the processor's application register set and two were used to exercise certain fictional units. Approximately 76% coverage of the application register set was obtained, and the fictional units tested were the ALU, FPU, and 1/0 buffers. Whenever an SEU was detected, the test conditions were recorded including the fluence (total number of particles until upset), device configuration, test program, ion, and LET.

The register set was tested by first writing a test pattern into all covered registers (either $FF, $00, $AA, or $55), and then continuously checking them for errors. After each iteration through the loop

a "heartbeat" character was displayed on the video monitor to indicate the program is still 'functioning. When a **miscompare** was detected, the radiation beam was stopped and the register and location of the upset was determined. The corrupted register was then written to main memory where the logic analyzer would read the register contents off the processor data pins. At this point, the correct pattern would be rewritten into the corrupted register and immediately read back. If a **latchup** condition was present the rewrite operation would fail and we would see the corrupted value again. If no **latchup** was detected, the beam was restarted and the program would continue checking all registers for errors again.

To test the ALU and 1/0 fictional units, the processor would perform a series of additions, subtractions, multiplications, and divisions. These results would be simultaneously written to a file for off-line analysis. A heartbeat character was also displayed on the video monitor to indicate which particular operation was being performed, as well as give a good health status check.

The FPU fictional test was done with a program obtained from one of the vendors, which performed stack checks, integrity checks, environment checks, as well as performed the same operations as in the ALU test.

C. Cyclotron Test Procedure

The Texas A&M University superconducting K-500 cyclotron facility was used to provide the heavy-ion radiation environment for the SEU testing. The ions used in the first test were Krypton (Kr), and Xenon (Xc), which provided linear energy transfer (LET) values of 25.1, and 43.1 (MeV cm^2/mg) respectively. A follow-up test is scheduled for November, 1996, which will use different ion beams.

The test board was mounted inside the cyclotron test chamber and positioned so that the radiation beam only irradiated the device under test (DUT). Translational movement over the X, Y, and Z axes plus angular control of the beam arrival angle was available during the test runs. All the data, test and power cables are routed into the chamber via a 50-pin D connector interface block. The test system was then closed and repressurized to an atmosphere of approximately 10-4 torr. When this was achieved, the ion beam was fine-tuned and focused before testing could commence. A shutter at the chamber entrance was used to control the precise time the beam was applied or removed to the DUT. The test procedure began each data run by rebooting the processor and loading a test program into memory. The test program would begin executing before the radiation beam was applied to the processor, and continue running until an error was detected or the program terminated.

III. TEST RESULTS

The first 486-DX4 radiation test was performed in September, 1996. Four microprocessors were tested producing a total of 135 data runs. Approximately 25 data runs were produced for ***each*** device. The test sequence was to first run all the test programs with the processor's internal L1 cache enabled, then with L1 disabled. This procedure was repeated for all four test devices and then again for each beam. The first ion beam used was Xe followed by Kr.

Eight distinct error modes were observed and were categorized into the following classes:

(1) **SEU** errors explicitly detected by test software: found only in the application registers
(2) data errors: most likely due to instruction or data upsets in the internal cache
(3) program-flow anomalies: test program experienced abnormal control flow
(4) program hangs: test program stopped working (i.e., program crash)
(5) floating-point unit failures: failures explicitly detected by the FPU test program
(6) system errors: upsets in a system-level register or instruction
(7) processor reboot: upsets that caused the processor to initiate a reboot
(8) **latchup**: detected by the power supply reaching its current limit value during an error

A. Data Discussion

Analysis indicates that ***program-hangs are*** the primary error mode for Vendor A, occurring 53.2% of the time. The cause for this was due to an illegal operation performed by the processor; such as would happen if a segment register, instruction pointer, or memory operand was upset. ***SEU-errors*** and ***program-flow-errors*** were the next two frequent error modes, occurring 17.8% and 9.3%, respectively. SEU-errors were detected in the register test by failing the test pattern compare. Program-flow-errors were observed when incorrect output messages or ascii "garbage" were displayed on the screen. The origin for these errors were probably the same as for program-hangs, but the location was such that the program could keep running. ***System-errors*** were observed only 6.5%, and were evidenced by error messages such as "internal stack overflow", "memory allocation error", or "no ROM basic." The *FPU_fail* errors was observed 1.9%, and could only be detected when running the FPU test program. Other errors were observed when running this test, primarily program-hangs and data errors. ***Reboot*** errors did not occur very often (1.90/0), and were usually seen in conjunction with another error. Only one ***latchup*** error was detected for Vendor A, and this occurred when running Xe (LET=43. 1 MeV cm^2/mg), thus Vendor A's latchup threshold is very near this LET. Figure 2 gives the upset cross-section of Vendor A for each of the two ion beams used. This illustration is a measure of the error rate for the device, and details of its calculation are given below in section B.

The primary error mode for Vendor B was *latchup*, occurring 100% of the time with Xe, and 47% of the time with Kr. No valid cross-section data was obtained using Xe, because the device would instantaneously latchup as soon as the beam was applied. From this testing, Vendor B was very susceptible to latchup and has a much lower threshold than compared to Vendor A. No destructive latchup was observed. Vendor B's chip is fabricated using a smaller feature size than Vendor A, so this may indicate that smaller feature sizes are directly proportional to higher upset rates, but further testing is required for confirmation. With the exception of latchups, program-hangs were also the most frequent error mode observed (36.80/0). The other errors seen in Vendor B were: SEU-errors, program-flow-errors, FPU-fail, and system-errors, each with a frequency of approximately 2.6°/0. Figure 3 shows the upset cross-section of the two ion beams for Vendor B.

The upset rate for ***all*** devices was found to be ***higher*** when the processor's internal L1 cache was enabled. This was expected since the test program and data are stored onboard the processor, and thus more vulnerable to radiation upsets. Figure 4 shows the processor's performance with the cache enabled versus disabled for Vendor A. Note that there is about an order of magnitude difference between these two operating modes, which is a significant difference in expected upset rate. Of course, with the cache disabled you lose performance, thus a trade-off must be made.

The data suggests that the testing performed was with the processor near the ***saturation level. This*** point occurs where any further increase in the linear energy transfer (LET) produces no increase in the cross-section upset rate. To conclude this 486-DX4 testing, it is very important that the initial part of the curve (x-intercept) be determined. This region is called the ***critical LET,*** and is the threshold at which the first occurrence of radiation-induced errors are observed. The goal of the follow-up radiation test in November, will be to find the critical LET of the 486-DX4.

B. Error Rate Calculation

To predict the error rate that can be expected in a radiation environment, an upset cross-section must be found by summing the total fluence for each unique error mode per device. In other words, the ***total*** number of particles were accumulated up to the first occurrence of ***each*** distinct error mode per device. The cross-section is then computed as:

$$\text{cross-section} = \frac{\text{\#of failures}}{\text{total fluence to error}}$$

where the # of failures is one for each upset. For subsequent errors, the total fluence is just the number of particles ***between*** each error for a particular error mode. A plot of this data versus the LET of each data point represents the capture upset cross-section that can be used to predict the susceptibility of the device to radiation. In Figures 2-4, the LET values for each beam have been expanded to allow easier interpretation of each error mode. Note the widely scattered values that cover several orders of magnitude. The reasoning for this, it is believed, is since the 486-DX4 is such a complex device, it is assumed to have many different failure modes that only manifested themselves in the eight error modes detected in this testing.

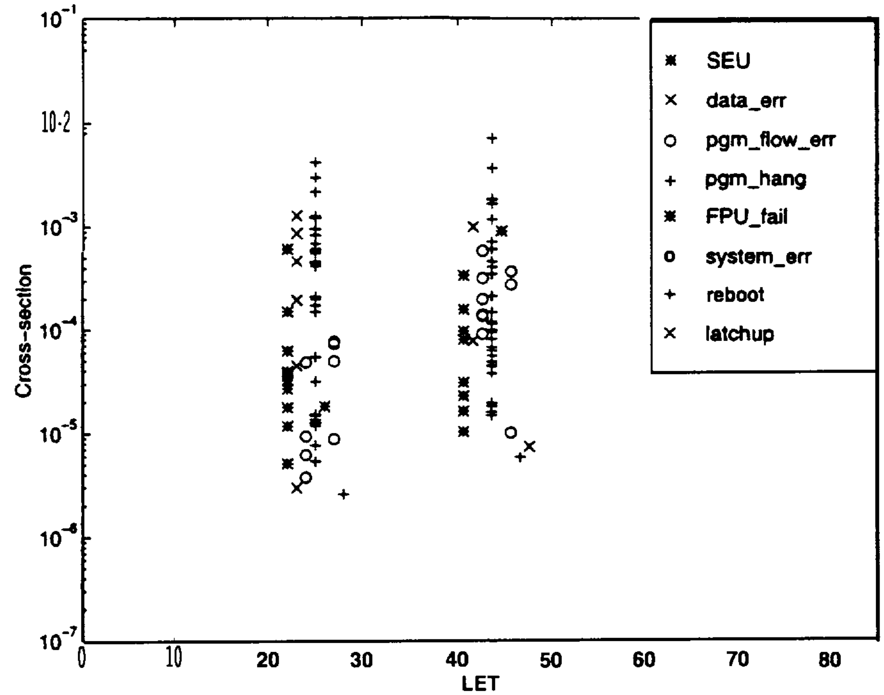


Figure 2: 486-DX4 Upset Cross-section vs. LET for Vendor A

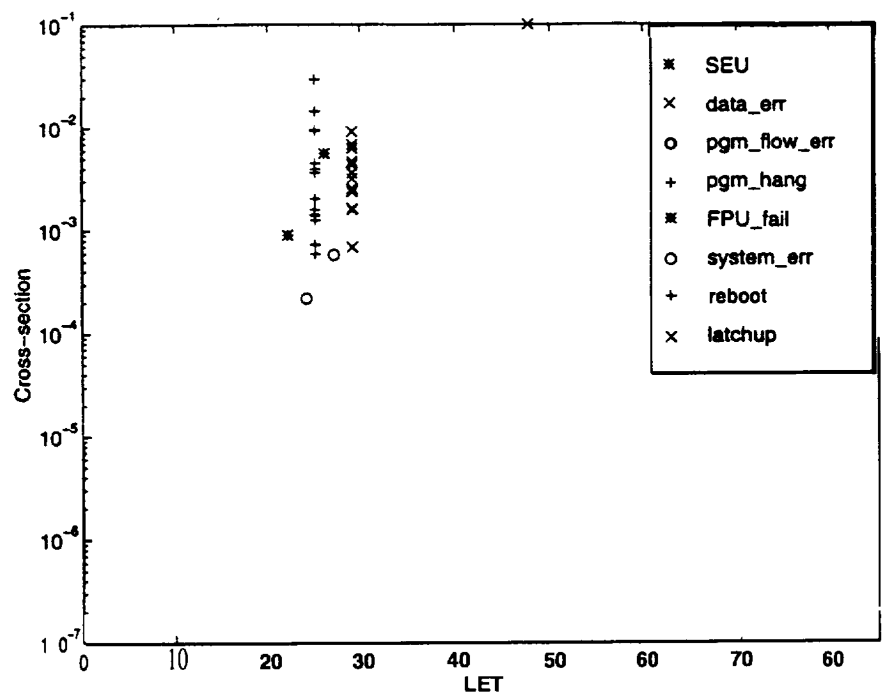


Figure 3: 486-DX4 Upset Cross-section vs. LET for Vendor B

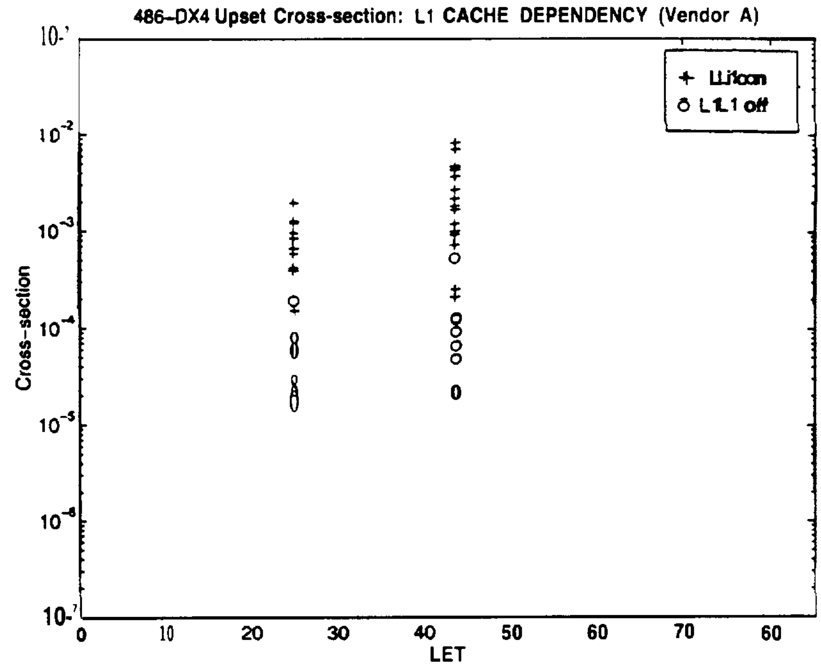


Figure 4: Upset-cross section illustrating Cache performance (Vendor A)

IV. CONCLUSIONS

From this testing, results indicate that the 486-DX4 is more susceptible to upsets with the internal L 1 cache enabled versus disabled. Results also show that Vendor B has a much lower latchup threshold than Vendor A. However, with the amount of test samples taken, and the varied degree in the upset cross-section, more testing should be done to corroborate these findings. The critical LET threshold will also need to be experimentally determined before a complete SEU analysis of the 486-DX4 can be made.

In summary, we believe this research is valuable to the future of digital devices that are used in space. As highly-reliable systems become more complex, traditional design evaluation & validation techniques that rely on experience and prior knowledge become impractical. It is therefore imperative to obtain accurate upset rates and behavior data at both the component level and the system level. Using a cyclotron as a fault-injection source provides a realistic means of simulating the space environment, and at a fraction of the cost of an actual spaceflight.

V. REFERENCES

[1] F.W. Sexton, "Single Event Upset in Space Environments," a White Paper, Div 2147, Radiation Technology and Assurance Division, Sandia National Labs, Albuquerque, NM

[2] D. Nichols, W. Price, C. Malone, "A Guideline for Heavy Ion Radiation Testing for Single Event Upset (SEU)," JPL Publication 84-22, Jet Propulsion Laboratory, Pasadena, CA, January 1, 1984

[3] P. O'Neill, "Space Shuttle Single Event Upsets Experience," Orbiter& GFE Project Office, NASA-Johnson Space Center, Houston, TX, March 10, 1994

[4] R. Rasmussen, "Spacecraft Electronics Design for Radiation Tolerance," Proceedings of the IEEE, Vol 76, No 11, pp. 1527-1537, November 1988

[5] R. Watson, H. Schwartz, and D. Nichols, "Test Report for Single Event Effects - 80386DX Microprocessor," Technical Report, let Propulsion Laboratory, California Institute of Technology, Pasadena, CA., February, 1993

URC97076

// ? ?
? ? / 12

How the Theory of Computing Can Help in Space Exploration

Vladik Kreinovich and Luc Longpré

Center for Theoretical Research and its
Applications in Computer Science (TRACS)
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968
email {vladik,longpre}@cs.utep.edu

Abstract

The opening of the NASA Pan American Center for Environmental and Earth Sciences (PACES) at the University of Texas at El Paso made it possible to organize the student Center for Theoretical Research and its Applications in Computer Science (TRACS).

In this abstract, we briefly describe the main NASA-related research directions of the TRACS center, and give an overview of the preliminary results of student research.

1 Preamble

The opening of the NASA Pan American Center for Environmental and Earth Sciences (PACES) at the University of Texas at El Paso made it possible to organize the student Center for Theoretical Research and its Applications in Computer Science (TRACS).

This center was one of the sponsors of the student regional conference in Computational Sciences SC-COSMIC (October 1996, El Paso, TX) [7], and it was instrumental in bringing the 1997 Annual ACM Symposium on Theory of Computing to El Paso.

The main emphasis of this center is not only on today's engineering problems, but also on the fundamental problems that can eventually be of use in different areas of practice, in particular, in the area of space exploration. In this abstract, we briefly describe the main NASA-related research direction of the TRACS center, and give an overview of its results.

2 Space-related computations: specific features, specific problems

How are computations and data processing related to space research different from computations in other application areas? There are many differences, regarding the *input,* the *computational resources* available for these computations, and the desired *result.*

Input. Many on-Earth computations deal with the reasonably well known areas, for which a lot of information already exists. Most space missions are, to the large extent, missions into the unknown. As a result, in computations related to space mission, there is usually a much smaller amount of a *priori* information.

The *amount of information* coming from the space missions is, usually, also *much smaller* than usual. There are several reasons for that; the three main reasons are:

- first, each additional sensor on-board costs a lot to launch and to maintain, and therefore, the designers try to keep the total number of sensors reasonably low;
- second, the more accurate and more reliable sensors are usually more complicated, weigh more, require more energy and other resources, and are usually less robust in the sense that may require too many

additional resources to protect them against the often hostile space environment; therefore, the sensors that are actually launched are often not the most accurate and the most reliable;

- finally, each additional bit of information has to be collected and transferred to the Earth, creating an additional burden on the spaceship communication system.

Computational resources. Processors launched into space (or placed on board of a planetary rover) must take as small amount of space, energy, etc., as possible. As a result, we must be able to perform all necessary computations on the available *limited* computational resources.

Desired result. Since space missions are usually very expensive, every potential error is very costly, and in manned space missions, errors are simply inadmissible. So, space-related computations must be 100% *reliable.*

3 Due to these specific features, problems of space-related computations are, in general, computationally intractable

Interval computations. First of all, in space-related computations, we need the *guaranteed* results. Together with the fact that the input comes from not 100% accurate sensors, this means that we have to estimate the accuracy of the results of data processing, while in traditional computing, accuracy is usually not an issue.

This bring up the importance of computations that take this inaccuracy into consideration and lead to guaranteed results. Due to the fact that we want not simply a *numerical estimate* of the desired quantity, but an *interval* within which the actual values of this quantity is *guaranteed* to lie, the corresponding area of theory of computing is called *interval computations.*

The problems become computationally intractable. In interval computations, there are important theoretical results, but it turns out that if we take the inaccuracies of the input data into consideration and require the results to be 100% reliable, then even the simplest tasks such as solving a liner system become, in general, computationally intractable (NP-hard) [10, 1].

What can we do?

4 How can we solve computationally intractable problems?

First approach: finding classes of solvable problems. The fact that the problem is computationally intractable means that there is no general algorithm for solving these problems. Therefore, a natural approach is to find *classes* of problems for which there are feasible algorithms. In this, theory of computing, with its large experience of designing new algorithms, can be of great help.

In particular,

- for general interval data processing, theory-motivated algorithms are presented in [11];
- for a specific problem of chemical identification, a problem that is very important for space exploration, algorithms based on theoretical ideas are presented in [4, 3].

Second approach: using human intelligence. In many real-life situations, the problem that we must solve does not belong to any of the classes for which feasible algorithms are known. In some of these cases, we know that human experts can often solve these problems really well. How can we describe the experience of these experts in such a way that the computer will be able to understand this knowledge and use it in data processing?

Here, there are three types of theoretical problems:

- is a certain type of formalism *sufficient* for a certain class of problems?
- what is a computational *complexity* of using this formalism?
- how can we make the resulting computations faster?

In TRACS, all three types of problems are analyzed:

- *Sufficient.* In [12], it is shown that one of the logic programming formalisms is sufficient to represent reasonably general type of knowledge. In [13], a similar problem of sufficiently is analyzed for *robots.*
- *Complexity.* In [5], it is shown that even if *we* manage to describe all the knowledge in terms of an expert system, the computational problems related to *using* this expert system are, in general, computationally intractable.
- *Faster.* In [9], a fast algorithm is developed for an important *particular* class of *fuzzy data processing* algorithms.

Third approach: trying to make the existing computers faster. If the existing algorithms are too slow, and no expert knowledge is available, then we can try to speed up the computers.

The design of a faster computer usually involves lots of engineering problems and heuristic methods, but, as it is shown in [8], theory of computing, with its experience of evaluating the computational abilities and computational times of different hypothetic computational devices, can definitely help in selecting the most promising design of a real-life computer.

Fourth approach: looking for radically new ways of computing. If the existing computers cannot solve our problems, then maybe some radically new ways of computing will be more helpful. In particular, one way to speed up computations is to miniaturize the computers.

At TRACS, we have looked into the potential abilities of *chemical computers [6]* that operation on the level of molecules and *quantum computers [2]* that operate on the level of elementary particles. Baaed on the results of this analysis, both approaches seem to be very promising.

Acknowledgments. This work was partially supported by NSF Grants No. CCR-9211174 and EEC-9322370, and by the NASA Pan American Center for Environmental and Earth Studies (PACES). The authors are thankful to Ann Gates and Scott St arks for their help and encouragement.

References

[1] G. Alefeld, M. Koshelev, and G. Mayer, "Fixed Future and Uncertain Past: Theorems Explain Why It Is Often More Difficult To Reconstruct the Past Than to Predict the Future", *These Proceedings.*

[2] A. Beltran, V. Kreinovich, and L. Longpré, *QFT + NP = P: Quantum Field Theory (QFT): A Possible Way of Solving NP- Complete Problems* in *Polynomial Time,* University of Texas at El Paso, Department of Computer Science, Technical Report UTEP-CS-96-45, November 1996; LaTeX file available as **ftp:** //es. **utep.edu/pub/report s/tr96-45.** t **ex**.

[3] A. Beltran and J. Salvador, "The Ulam Index: Methods of Theoretical Computer Science Help in Identifying Chemical Substances", *These Proceedings.*

[4] A. Beltran and J. M. Salvador, "The Ulam index", *Abstracts of the Second SC-COSMIC Conference in Computational Sciences, October 25-.27, El Paso, TX,* Rice University Center for Research on Parallel Computations and University of Texas at El Paso, 1996, p. 6.

[5] L. Chee, *Computing the Value of a Boolean expression with intervals is NP-hard*, Master Thesis, Department of Computer Science, University of Texas at El Paso, 1996.

[6] B. Cloteaux, On *the Computational Power of Using Chemical Reactions,* Master Thesis, Department of Computer Science, University of Texas at El Paso, 1996.

[7] *Abstracts of the Second SC-COSMIC Conference in Computational Sciences, October 25–27, El Paso, TX,* Rice University Center for Research on Parallel Computations and University of Texas at El Paso, 1996.

[8] M. Hampton, "How difficult is it to add 1? A pedagogical example of how theory of computing maybe useful", *These Proceedings.*

[9] M. Hampton and O. Kosheleva, "Fast Fuzzy Arithmetic Operations", *These Proceedings.*

[10] P. Kahl, Solving *Narrow-Interval Linear Equation Systems Is NP-Hard*, Master Thesis, Department of Computer Science, University of Texas at El Paso, 1996.

[11] M. Koshelev and P. Taillibert, "Optimal Approximation of Quadratic Interval Functions", *These Proceedings.*

[12] O. Kosheleva, "An Arbitrary First Order Theory Can Be Represented by a Logic Program: a Theorem", *These Proceedings.*

[13] M. Nogueira, "What Can Robots Do? Towards Theoretical Analysis", *These Proceedings,*

Solar Cycle Variation of Atmospheric Nitric Oxide

C. Krishna Kumar *, A. P. Batra*, and L. Klein *'
Department of Physics& Astronomy
Howard University, Washington, DC 20059

Arun Batra
Thomas Jefferson High School for Science and Technology
Alexandria, VA 22312

Introduction

Concentration of nitric oxide in the upper atmosphere has been measured with the Halogen Occultation Experiment (HALOE) instrument in the Upper Atmosphere Research Satellite (UARS). The database from this experiment contains NO volume mixing ratios at altitudes up to about 145 km at intervals of 5 km and covers up to 80 degrees in latitude on either side of the equator. The time period covered is from October 91 to February 1996. This period started with high solar activity, F10.7 = 250 down to 75 now. In this report observations in the 30-50 km and 80-140 km ranges will be analyzed both for the equatorial zone 30 N to 30 S and for the region poleward of 50 degrees. The purpose of this investigation is to study the solar cycle variation of nitric oxide in the atmosphere.

Earlier experiments with the Solar Mesosphere Explorer (SME) satellite showed that the variation in the NO density at 110 km correlated with both long -term variation in solar activity for the time period 1982-86, a period of high and decreasing solar activity [Barth et al., 1988; Barth, 1992]. Barth claims that the NO concentration is correlated with solar rotation also but presents no quantitative evidence. Furthermore the UV measurements of NO are known to have systematic errors because optical depth effects were neglected. Kuze and Ogawa (1988, and references therein) claim evidence for variation of nitric oxide with solar cycle, based on rocket measurements. The HALOE data are the longest daily record, to date, of NO measurements with the same instrument. They provide global coverage also.

Observations

From October 91 to date HALOE instrument has made several (up to 30) observations each day of NO. Fifteen of these are at sunrise with the tangent point in one hemisphere and the other ten at sunset in the other hemisphere. The volume mixing ratio, pressure, temperature, and NO data uncertainty have been derived from the UARS data as a function of altitude and latitude. These data are available from the NASA/LARC data center. From these data, the vertical column densities of NO and their associated uncertainties were derived. The column densities were separated into two altitude ranges, 30-50 km and 80-140 km. For each of these altitude ranges, daily averages were obtained for two latitude regions - equatorial zone from 30 N to 30 S and merged polar regions with latitudes poleward of 50 N and 50 S. Further, sunrise and sunset data were separately analyzed. The variation of latitude of sunset or sunrise is about 1° each day. The uncertainty of the mean column density was estimated from that of the individual measurements, assuming that they were not correlated. The uncertainties in the daily averages of vertical column densities are less than 20%. This report used version 17 of the retrievals. Version 18 became available even as this report is being written. It is believed

[1] Center for Study of Terrestrial and Extra-terrestrial Atmospheres

that version 18 values may differ from the version 17 values by no more than 15 %. Therefore changes in NO column densities by more than 15 % are robust.

Results

Figures 1-4 shows plots of daily averages of NO vertical column densities obtained during the last five years. The data analyzed are from October 1991 to February 1996. The conclusions one can draw about solar cycle variation from an examination of the figures are: (i) there is no variation with solar cycle at low altitudes (30-50 km), (ii) the variation at high altitudes (80-140 km), from high solar activity (October 199 1) to the its low value (February 1996) is a factor 3 at low latitudes and about the same in the polar regions. The secular variation in the polar regions is masked to some extent by the very large scatter in the sunset results. Another conclusion is that the scatter in the column densities is larger in the polar regions than in the equatorial belt. There is a difference between the sunrise and sunset values in all the regions investigated except the high altitude equatorial values. Atmospheric chemistry models predict little or no difference between sunset and sunrise densities of NO at high altitudes (> 80 km) and the results here are in general agreement with that. But at low altitudes the models predict differences of less than 20 % whereas the observed value are 50% at least. Finally a close examination of the daily averages show a tendency for the value to increase as one goes toward the equator (slopes in the points) . This may bean instrumental effect in version 17 which is expected to be removed from version 18 data.

Future work

It is planned to estimate quantitatively the coefficients of correlation of each set of column densities with the indices of solar activity, F 10.7, solar x-ray flux from YOHKOH satellite, and the geomagnetic indices Ap and Kp. In addition more detailed comparison with models of atmospheric chemistry will be made to understand the sunrise-sunset differences.

References

Barth, C. A., W. K. Tobiska, D. E. Siskind, and D. D. Cleary, Solar-terrestrial coupling: Low -latitude thermospheric nitric oxide, (1988), *Geophys. Res. Letters, 15,92-94.*

Barth, C. A., Nitric oxide in the lower thermosphere, (1992), *Planet. Space Sci., 40, 315-336.*

Kuze, A.. and T. Ogawa, Solar cycle variation of thermospheric NO (1988): A model sensitivity study, *J. Geomag. Geoelectr., 40, 1053-1065.*

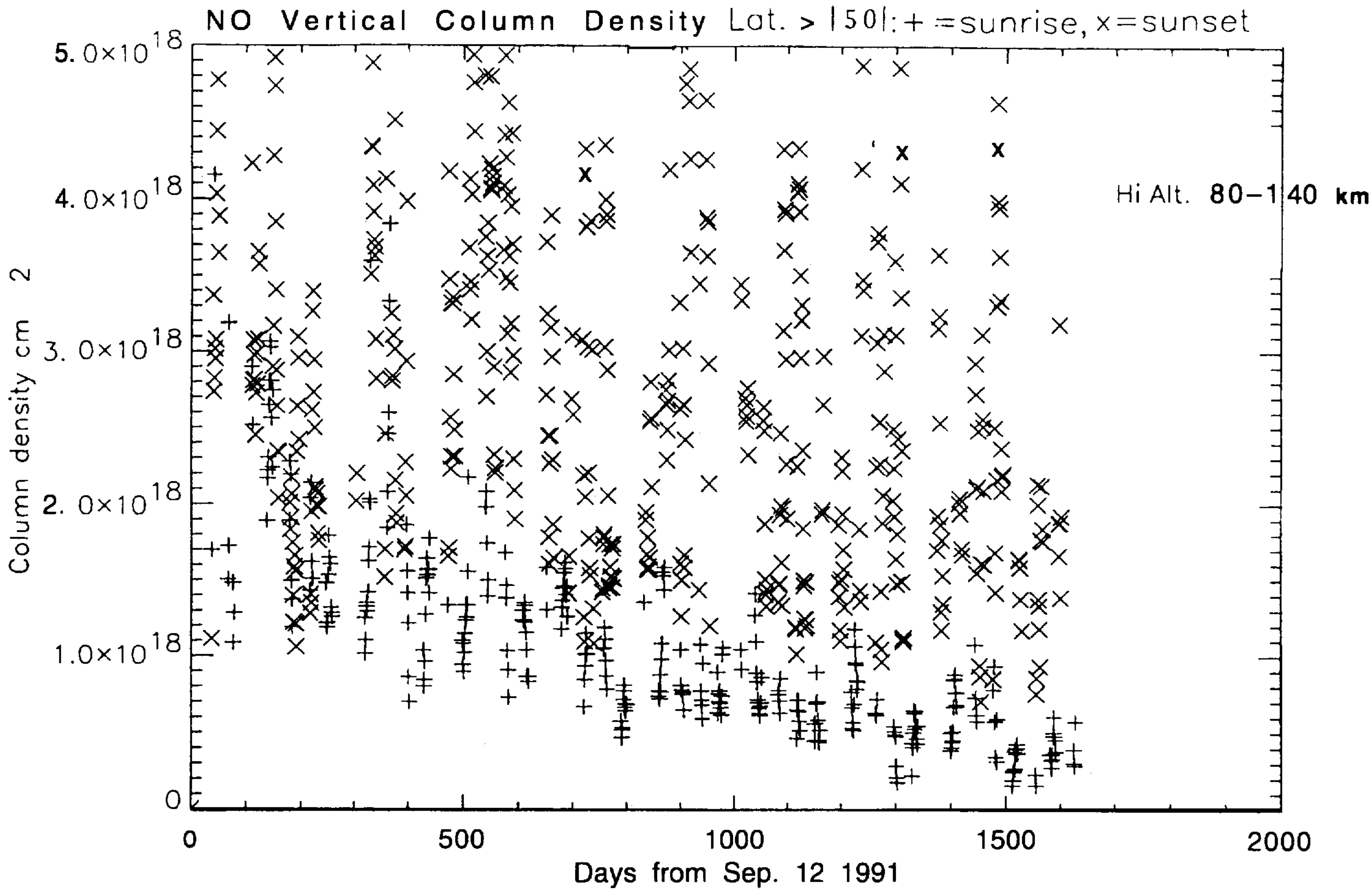


Fig. 1. Daily **average of NO vertical column density versus day from** September 12, 1991 **for combined latitudes poleward of 50 N and 50 S and for altitudes 80-140 km.**

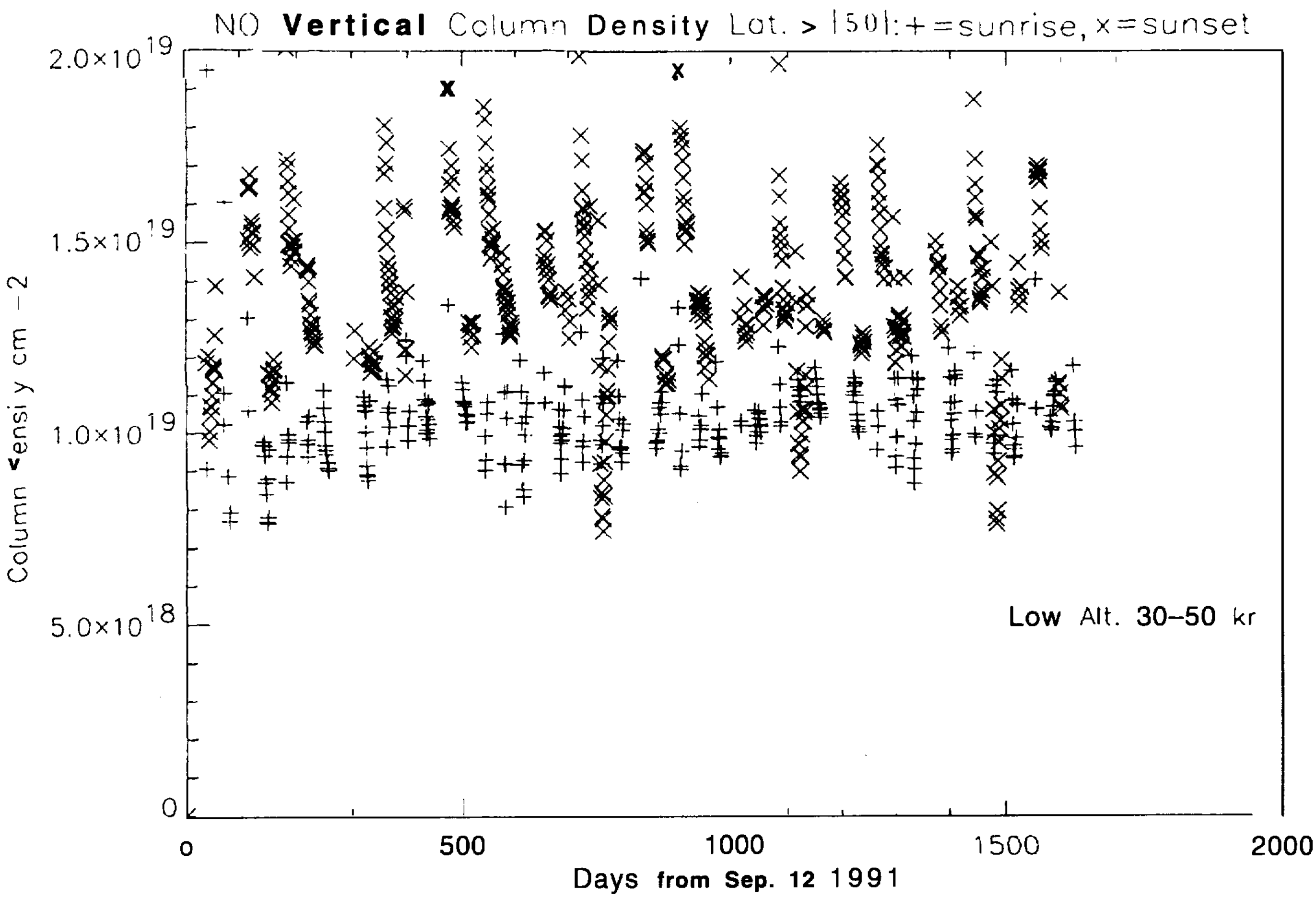


Fig. 2. Daily **average of NO vertical column density versus day from September 12, 1991 for combined latitudes poleward of 50 N and 50 S and for altitudes 30-50 km.**

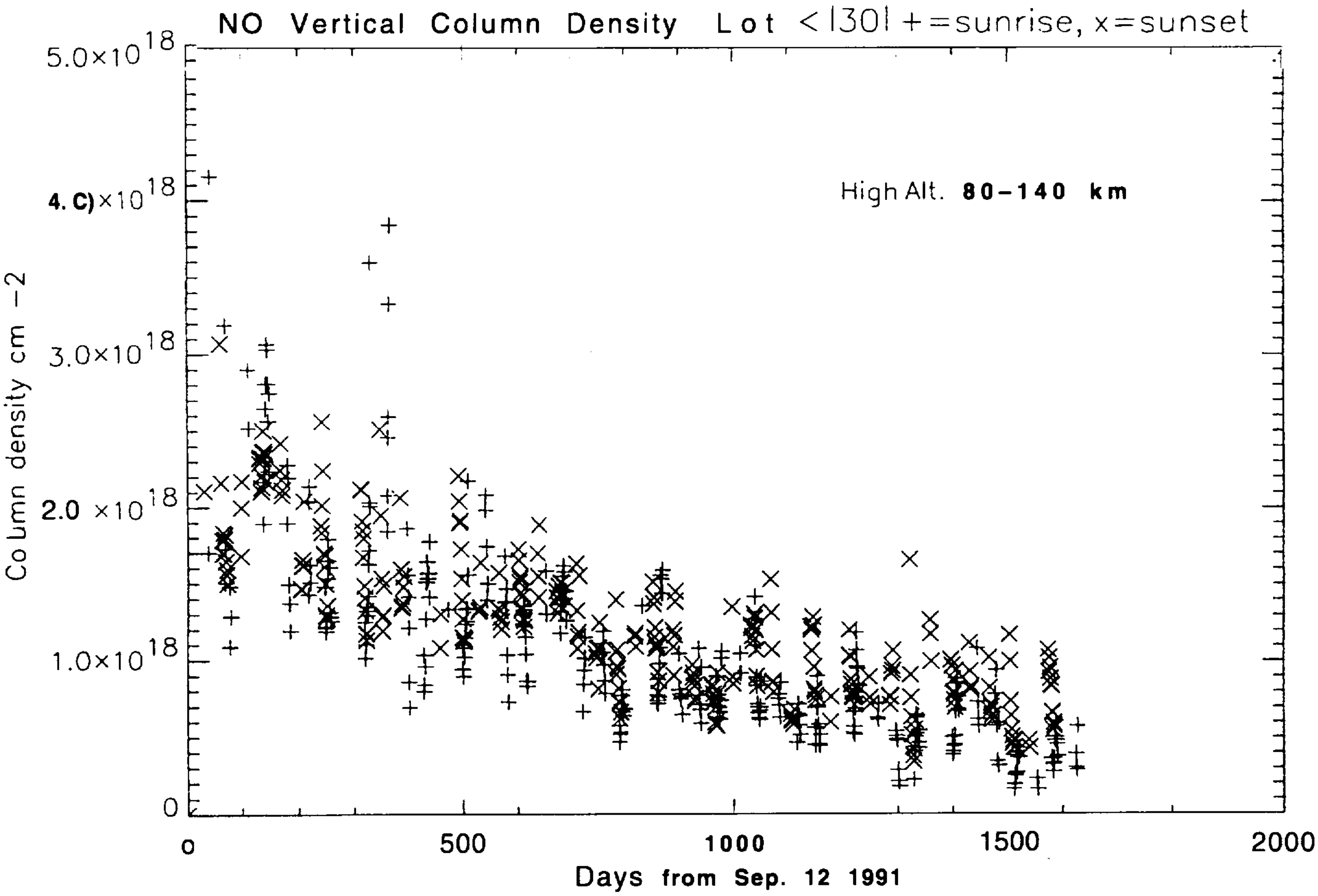


Fig. 3. **Daily average of NO vertical column density versus day from September 12, 1991 for equatorial zone between 30 N and 30 S and for altitudes 80-140 km.**

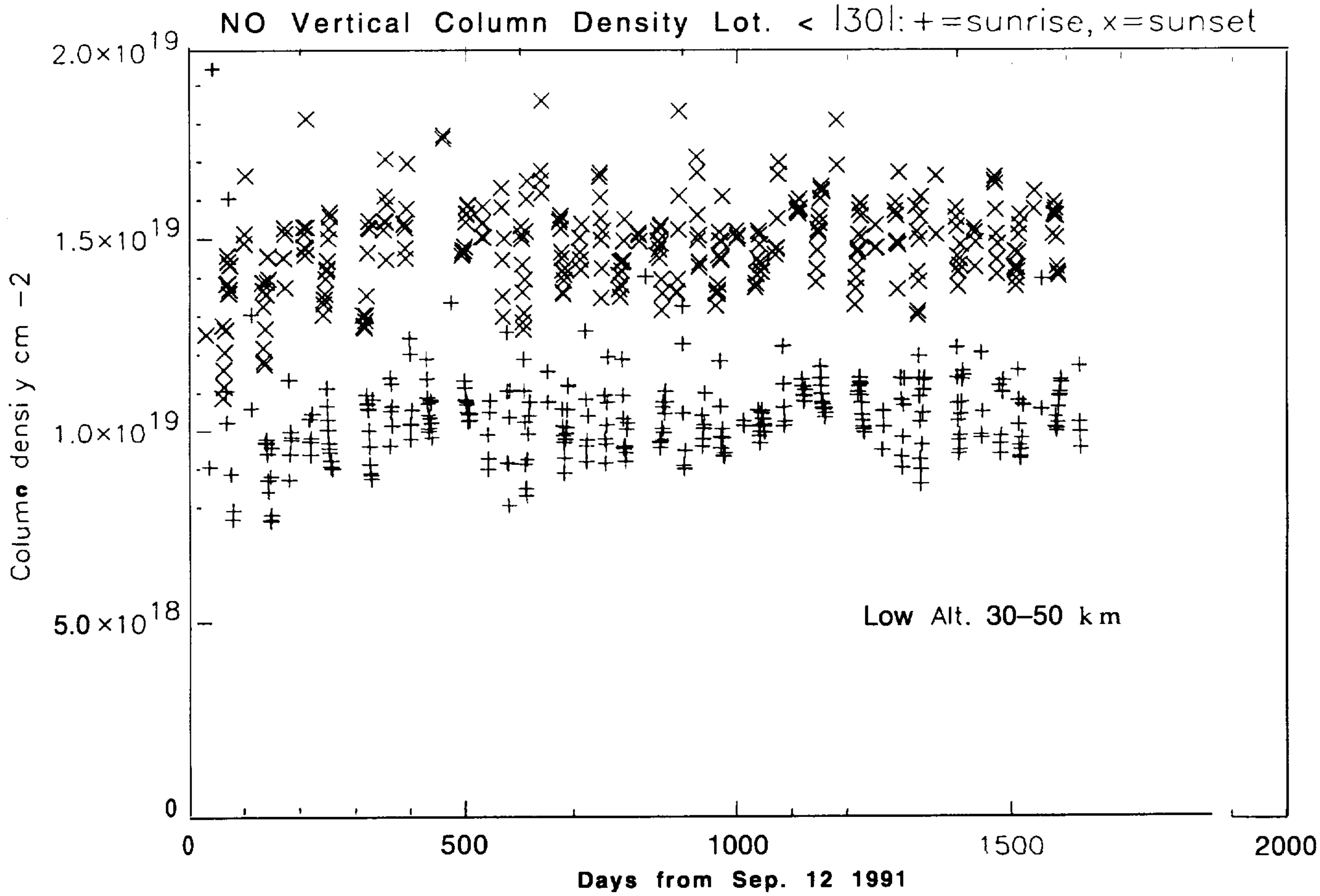


Fig. 4. **Daily average of NO vertical column density versus day from September 12, 1991 for equatorial zone between 30 N and 30 S and for altitudes 30-50 km.**

Neural Network Based Sensory Fusion for Landmark Detection[2]

Kishan -K. Kumbla and Mohammad-R. Akbarzadeh-T.
Center for Autonomous Control Engineering
Electrical and Computer Engineering Department
University of New Mexico, NM 87131
E-mail : kkumbla@eece.unm.edu

Abstract

NASA is planning to send numerous unmanned planetary missions to explore the space. This requires autonomous robotic vehicles which can navigate in an unstructured, unknown, and uncertain environment. Landmark based navigation is a new area of research which differs from the traditional goal-oriented navigation, where a mobile robot starts from an initial point and reaches a destination in accordance with a pre-planned path. The landmark based navigation has the advantage of allowing the robot to find its way without communication with the mission control station and without exact knowledge of its coordinates. Current algorithms based on landmark navigation however pose several constraints. First, they require large memories to store the images. Second, the task of comparing the images using traditional methods is computationally intensive and consequently real-time implementation is difficult. The method proposed here consists of three stages, First stage utilizes a heuristic-based algorithm to identify significant objects. The second stage utilizes a neural network (NN) to efficiently classify images of the identified objects. The third stage combines distance information with the classification results of neural networks for efficient and intelligent navigation.

1. Introduction

Mobile robot navigation based on landmark detection is an emerging area of research and is particularly significant for autonomous vehicles used to explore planetary surfaces. This becomes even more important if the mobile robot has to navigate in an unstructured, unknown and obstacle ridden environment. The additional intelligence incorporated in the navigational algorithm results in better path planning, avoidance of limit cycle operation (going to the same place or looping around a particular place) and hence lower power consumption. Traditional goal-oriented algorithms, which rely on accurate position and orientation determination, depend on a pre-planned path and consequently are unsuitable for this application. Implementation of landmark navigation using classical image processing techniques, however, poses several constraints. Recording hundreds of landmark images requires vast amounts of memory and storage space. Furthermore processing these images to identify landmarks is a computationally intensive task.

Neural networks (NN) have shown considerable promise in classifying images and require less memory. The classification of significant images as landmarks is performed in the learning mode of the neural network when it updates its neuron interconnection weights. Ultrasonic sensory data of the distances of the *significant objects* in a particular landmark is recorded. This serves as reaffirming information when the robot is trying to detect the already learnt landmarks. Detecting relevant objects and images is another problem. This we propose to approach using an existing

[2]This work was supported in part by NASA contract # NCCW-0087.

U.S. patented technology developed for *Fuzzy Control of Video Printer.* This technology uses the fuzzy logic based algorithm to detect relevant objects to improve the quality of the color. This entire method of using neural network with ultrasonic sensory data warrants limited memory requirements and can be implemented in real-time as it is computationally less intensive. Incorporating the emerging technology of neural network for detecting landmarks is a new and innovative method with practical application. Ultrasonic sensors can be replaced by laser range finders or any other distance sensors if ultrasonic sensors do not work on some planetary surfaces (e.g. devoid of atmosphere),

2. Related Work

In 1996, NASA will launch the Mars Pathfinder spacecraft, which will carry an 11-kg rover to explore the immediate vicinity of the lander [1]. At present the rover navigation is divided into four major functions: goal destination, rover localization, hazard detection and path selection. There is significant room for improvement of the navigational algorithm by adding extra degrees of intelligence. This can be done by incorporating a landmark based navigation algorithm. Using fuzzy logic and neural network based methods, the efficiency, robustness, adaptability and accuracy of navigation can be vastly enhanced. This will allow autonomous operation without communication from the lander and reduction of power consumption by selecting shorter paths by avoiding limit cycle operation of the rover.

There are several approaches to local positioning of mobile robots using landmark detection. These approaches generally apply only to structured environments or pre-defined landmarks. Nishizawa, et. al. [2] describe a local positioning scheme using landmark detection where the landmarks are reflectors placed sparsely in the ceiling of the robot's work space. They then utilize the Maximum Likelihood Estimation technique to eliminate the gradual error accumulation caused by the integration of the odometer measurement. Fukuda et. al. [3] uses a fuzzy template matching to recognize edges of air conditioning outlets in the ceiling to be used as landmarks. A NN then determines if the object is an air conditioning outlet. This information is then compared with the map of the robot's work space in order to position the robot. In [4], Jung proposes a robot vision system modeled after human's, where peripheral vision and central vision are processed separately for a more efficient detection. Luo and Potlapalli [5] utilize fractals to recognize landmarks in an outdoor environment. Since fractals are inherently scale invariant, the algorithm is robust to changes in light intensity and viewing direction of the landmark. Krotkov and Hoffman [6] developed a quantitative model of surface geometry in order for a safe, power efficient locomotion over the natural, rugged terrain. Ricotta and Liotta [7] describe a real time visual landmark tracking system for mobile robots. They discuss image processing techniques to track "relevant" objects in the robot surroundings and keep desirable objects within the robot's field of view. In [8-10], other approaches to landmark detection in structured environments are discussed.

Fuzzy logic has already been applied to data fusion. Murphy [11] describes a sensory fusion effects architecture for robot navigation. Also, he discusses the utility of biological and cognitive insights for sensory fusion. Abidi, et. al. [12,13] use fuzzy logic to integrate several sensory data and to enhance the recognition capability of an autonomous system by yielding meaningful information which are otherwise unavailable or difficult to acquire by a single sensory modality. Zhang et. al. [14] applied fuzzy logic for integration of deliberative and reactive strategies, where

programming at the task level in a partially known environment is divided into two consecutive steps: sub-goal planning and sub-goal guided plan execution. In [15], Song et. al. describe an integration routine for an ultrasonic sensor and a CCD camera. They utilize an extended discrete Kalman filter to fuse raw sensory data and to provide a more reliable representation for environment perception.

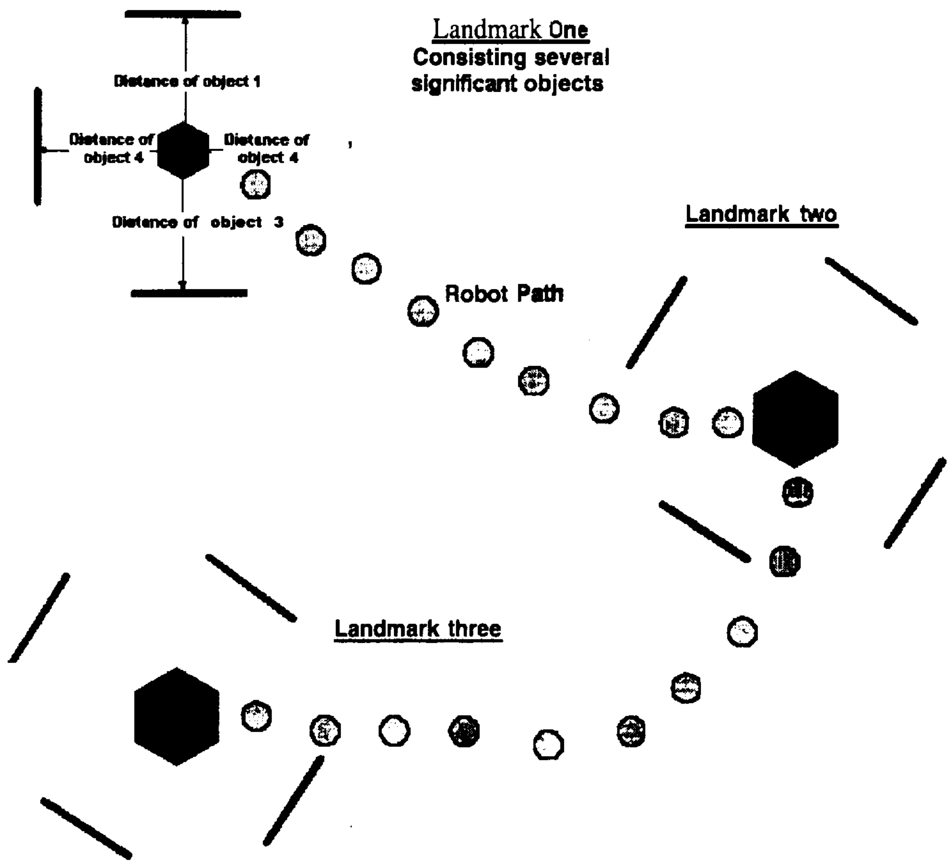


Figure 1: Landmark, significant object and robot path

3. Fusion of information for landmark detection

3.1 Significant Objects and Landmarks

The primary function in this method of landmark detection is to identify relevant or significant objects. The digital image frame from the camera is used as an input to a fuzzy logic based object identification algorithm. This algorithm decides whether the image can be classified as a distinguishable significant objects. This information depends on the environment in which the mobile robot is supposed to navigate. Such significant objects are represented by a sharp contrast or color variation in the image. In the actual settings, these landmarks need to be reaffirmed by other sensory information such as distance measurements. This can be done by using ultrasonic sensors or laser range finders. This means the object can only be considered for a landmark only if the position sensor can determine the distance from the object. A landmark can be a single object or a collection of several significant objects. By having several objects the identity of the landmark can be made unique. A sensory fusion algorithm based on fuzzy rule-sets can be used to fuse the information on a particular landmark. Figure 1 shows a landmark consisting of several

objects. We propose the use of ultrasonic sensors for implementation of the landmark detection algorithm. Figure 1 describes the significant object, landmark and mobile robot path. As the robot is traversing the unknown path, it continuously monitors the terrain. It looks for significant objects which can be considered as landmark.

3.2 Classification of landmarks using Neural Network

A multi-layer perception Neural Network (NN) has been extensively used to classify images. Once the fuzzy logic algorithm detects a valid object, the NN is trained to recognize this a particular object. Several images associated with the landmark is trained to the NN's memory by a back-propagation training algorithm. After the NN has successively learnt a landmark the robot can look for another landmark. The NN is trained to recognize several landmarks while the robot is traversing a terrain. Although there is no set size for the NN for learning a particular set of landmarks, it is generally based on the size of training data; the NN size increases as the number of images to be classified increase. In this proposed method a multi-layer network with two hidden layers. The number of neurons in the input layer is determined by the number of pixels in the image. The output layer neurons would correspond to the maximum number of landmarks to be detected.

Even though the usage of neural networks allows for reduction of memory requirements in identification of landmarks compared to actual storage of images, a large number of landmarks require an increased network size. Therefore, the mobile robot's allotted onboard memory will dictate the network size and consequently the number of landmarks which can be detected.

Figure 2 shows the flowchart of the comprehensive landmark detection algorithm. A combination of visual image and distance information is used to detect a landmark. A fuzzy logic based algorithm identifies significant objects. At the same time the sonar measures the distance of the object. If the distance can be accurately measured then the neural network starts the learning algorithm and classifies the significant object as a landmark. For any particular landmark several nearby objects are classified and the distance and orientation information are measured. A landmark thus is determined by fusion of image information with the distance information by a fuzzy logic based fusion algorithm.

3.3 Hardware and Software

The facilities at NASA ACE will allow us to develop the software for simulations as well as for future experimental validation. The digital camera will be directly interfaced with a Pentium-based personal computer. We have access to several possible platforms of mobile robots which includes an in-house built mobile robot, *LOBOT, a Cybermotion's NAVMASTER*, and *Angelus Research's WHISKERS Intelligent Autonomous Vehicle.* These mobile robots are equipped with several sensors including ultrasonic sensors. Visual C^{++} will be used for basic software development. We will use an adaptive fuzzy logic control algorithm, *Dynamic Fuzzy,* which has already been developed. *Lab Windows* will provide a user friendly visual interface.

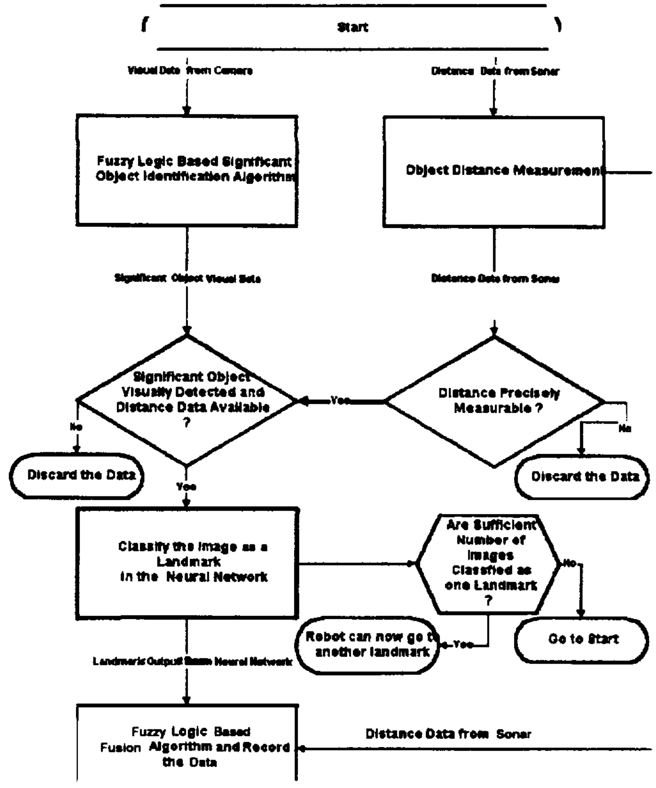


Figure 2: Flow Chart of the Landmark Detection Algorithm

4. Conclusions

This paper discusses a proposal to identify landmarks using a combination of fuzzy logic based significant object recognition algorithm and a neural network for classifying the landmarks. The landmarks are detected and defined by a combination of visual and distance information which are fused together for effective navigation. The proposed scheme will be implemented on real-time using the available mobile robot test-beds. An additional application of this method is related to robots on the factory floor. Using this method, the robots can identify parts in an assembly line to decide on a particular task. This eliminates the necessity of positioning the parts accurately which in turn allows a more flexible robotic assembly.

References

[1] L. Matthies, E. Gat, H. Ried, B. Wilcox, R. Volpe and T. Litwin, "Mars Microrover navigation: performance evaluation and enhancement", Autonomous Robots V 2 N4, 1995, Kluwer Academic Publishers Dordrecht, Nehterland, pp 291-311.

[2] T. Nishizawa, A. Ohya and S. Yuta, "Implementation of on-board position estimation for mobile robot - EKF based odometry and laser reflector landmarks detection", Proceedings - IEEE International Conference on Robotics and Automation, Part 1 (of 3), May 21-27, 1995, Nagoya, Japan.

[3] T. Fakuda, S. Ito, F. Arai, Y. Yokoyama,Y. Abe, K. and Tanaka Y. Tanaka, "Navigation system based on ceiling landmark recognition for autonomous mobile robot- landmark

detection based on fuzzy template matching, IEEE International Conference on Intelligent Robots and Systems Proceedings, Part 2 (of 3), August 5-9, 1995, Pittsburgh, USA.

[4] H. C. Jung, "Visual navigation for mobile robot using landmarks", Advanced Robotics, V 9 N 4, 1995 VSP Int. Sci. Publ. Zeist Netherlands, pp 429-442.

[5] R. Luo and H. Potlapalli, "Fractal based outdoor landmark recognition system for the navigation of a mobile robot", Proceedings- IEEE International Conference on Robotics and Automation, May 8-13,1994, San Diego, CA, USA.

[6] E. Krotkov and R. Hoffman, " Terrain mapping for a walking planetory rover", IEEE Transactions on Robotics and Automation, V 10 N 6, December 1994, pp 728-739.

[7] M. Ricotti and A. Liotta, "Real-time landmark detection for the mobile robot PARIDE', Proceedings of SPIE- The International Society for Optical Engineering Vision Applications in Industrial Inspection III February 8-9, 1995, San Jose CA, USA.

[8] A. Gilg, and G. Schmidt, "Landmark-oriented visual navigation of a mobile robot", IEEE Transactions on Industrial Electronics, V 41 N 4, August 1994, pp 392-397.

[9] S. Atiya and G. D. Hager, "Real-time vision-based localization", IEEE Transactions on Robotics and Automation, V 9 N 6, December 1993, pp 785-800.

[10] D. Kortenkamp, M. Huber, C. Congdon, S. Huffman, C.R. Clint, C.J. Cohen, F. Koss, U. Raschke and T.E. Weymouth, "Integrating obstacle avoidance, global path planning, visual cue detection and landmark triangulation in a mobile robot", Proceedings of SPIE- The International Society for Optical Engineering Mobile Robots VII, November 18-20, 1992, Boston, MA, USA.

[11] R.R. Murphy, "Biological and cognitive foundations of intelligent sensor fusion", IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans V 26 N1, January 1996.

[12] M.A. Abidi, M. Abdulghafour and T. Chandra, "Fusion of visual and range features using fuzzy logic", Control Engineering Practice, V 2 N 5, October 1994, Pergamon Press Ltd. England, pp 833-847.

[13] M. Abdulghafour, T. Chandra and M. Abidi, "Data fusion through fuzzy reasoning applied to segmentation of multisensory images", Proceedings of SPIE- The International Society for Optical Engineering Neural and Stochastic Methods in Image and Signal Processing, July 20-23, 1992, San Diego, CA, USA.

[14] J. Zhang, F. Wine and A, Knoll, "Modular design of fuzzy controller integrating deliberative and reactive strategies", Proceedings- IEEE International Conference on Robotics and Automation, Part 3 (of 4), April 22-28, 1996, Minneapolis, MN, USA.

[15] K. Song and W.-H. Tang, "Environment perception for a mobile robot using double ultrasonic sensors and a CCD camera", IEEE Transactions on Industrial Electronics, V 43 N3, June 1996, pp 372-379.

URC97079

Fuzzy Control of an Inverted Pendulum using Sensory Fusion and Hierarchy

57 / 63

Véronique Lacrose* and André Titli**

(*) LAAS du CNRS
7 Avenue du Colonel Roche, 31077 Toulouse Cedex - France
Tel. +(33) 0561336962- Fax. +(33) 0561336936
Email: lacrose@laas.fr

(**) LAAS du CNRS and INSA Toulouse, *France*
visiting NASA ACE center, University of New Mexico, USA
Tel. 5052770300- Fax. 5052773158
Email: titli@unm. edu

Abstract

A major problem of fuzzy control when handling complex and large-scale systems is that the number of rules grows exponentially with respect to the number of sensory input variables of the controller. This problem can be dealt with by application of rule-base reduction methods like sensory fusion and hierarchy. The above mentioned reduction methods were applied to a classical control problem: the inverted pendulum problem. Tuning of the scaling factors was realized by means of a gradient descent method. Satisfactory performance was achieved with a fuzzy logic controller of only 5 rules.

Keywords

Fuzzy control, Rule-base reduction, Self-tuning, Gradient descent method.

I. Introduction

If fuzzy control is applied to large-scale systems one of the major problems of this control method becomes clear, namely the number of rules grows exponentially with the number of input variables. Indeed, if we want to keep the completeness property of a rule-base, a conventional fuzzy controller with n input variables would result in $r = m^n$ rules, with m the number of Linguistic terms per input. It is necessary, for an easier implementation of a Fuzzy Logic Controller (FLC) and for real-time constraints, to reduce the number of rules. This problem can be dealt with by application of rule-base reduction methods.

In section II, sensory fusion and hierarchical approaches are briefly reviewed. In Section III, the above mentioned reduction methods were applied to the well-known inverted pendulum problem. Tuning of the scaling factors was done with the help of a gradient descent method. Section IV concludes the paper.

II. Rule-base reduction methods

In the last couple of years several rule-base reduction methods [1], [2], [3] have been proposed. In what follows the use of sensory fusion and hierarchical structures for rule-base reduction will be discussed. It will be shown that a combination of these both methods gives the most interesting results. In that case, the FLC consists of sub controllers that have only 2 inputs at the most and 1 output, which leads to an easier rule base writing.

A. Sensory fusion

The sensory fusion approach consists in combining sensor signals before injecting them as input for the FLC. These variables are often fused linearly (see Fig. 1).

It is assumed, in Fig. 1, that the input signals of the FLC are represented by $m = 5$ linguistic labels. a, b and c are positive parameters dictated by physical considerations, designers knowledge or experience.

Using sensory fusion, the reduction has a lower bound if all variables could, somehow, be fused. However, it is clear that all variables cannot be combined trivially. Every combination has to be reasoned and explained. In practice, only 2 variables are fused: generally the error and the change of the error.

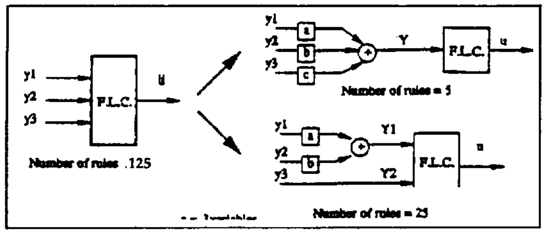


Fig. 1. Sensory fusion for n = 3

B. Hierarchical structures and fusion

The hierarchical fuzzy controller (see Fig. 2) was first introduced by Raju et al. [2]. B y this hierarchical structure, the number of rules will increase linearly (not exponentially) with the number n of system variables. Deciding where the variables are put into the hierarchy is an extremely important but difficult process, mostly based on knowledge of the system and sensitivity analysis. In practice, the variables are classified in accordance to their importance.

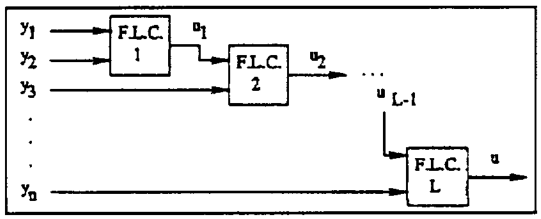


Fig. 2. The hierarchical fuzzy controller

Combining both methods discussed above [1] leads to the so-called 'Hierarchical and sensory fusion approach' (see Fig. 3). Here, the variables are simply combined first, as in Fig. 1, and are then organized into a hierarchical structure similar to that of Fig. 2.

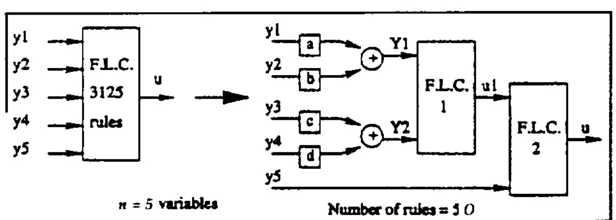


Fig. 3. Hierarchical and sensory fusion approach

In Fig. 4 the performances of the different approaches for rule reduction are expressed graphically.

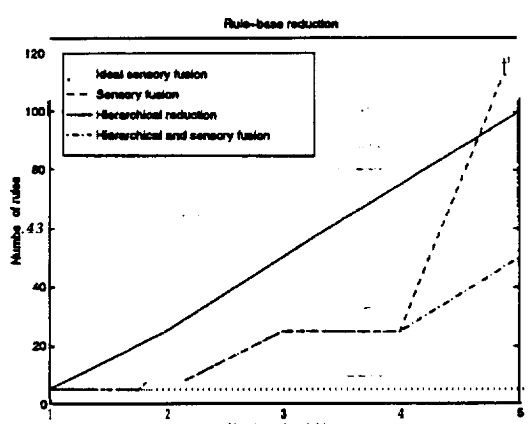


Fig. 4 Comparison of some rule-base reduction methods

One may note that ideal sensory fusion which could be obtained by fusing all variables, is impossible

in practice. Let us recall that the fusion of an error signal e and its derivative $\dot{e}$ is usually exceedingly successful.

III. Application to the inverted pendulum

In order to validate the rule-base reduction methods that have been proposed, these will be applied to the inverted pendulum (Fig. 5).

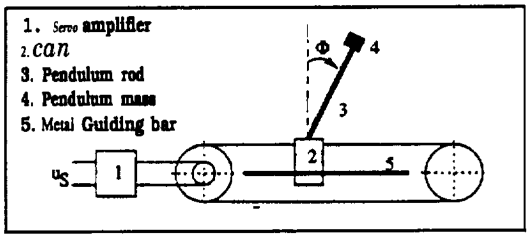


Fig. 5. Scheme of the inverted pendulum

The input variables of the fuzzy controller are the angle@ and the angular velocity $\Delta\Phi$, together with the error in position e given by $e = r - r_{ref}$ (where r is the position of the cart and r_{ref} the reference position) and the change of error Δe. The FLC's output variable is the control signal $u_S \epsilon [-10V, +10V]$.

The objectives are to control the pendulum in position and by the meantime to keep the pole balanced.

Tuning of the FLC scaling factors is done through the minimization of a quadratic criterion

$$J = \sum_{t=t\_begin}^{t=t\_end} \left(\frac{1}{2} q_{11} \cdot [r(t) - r_d(t)]^2 + \frac{1}{2} q_{22} \cdot \Phi(t)^2 \right)$$

by use of a gradient descent method.

r_d is a rough reference model for the position and is the result of a linear interpolation between four time points (see Fig. 6): the delay time (τ), the rise time (t_r at 10%), the establishment time (t_e at 2%) and the end time of the transitory period (to). The chosen time parameters for the position reference model are the following ones: $\tau = 0.36s$, $t_r = 2.5s, t_e = 3.5s$ and $t_p = 5s$.

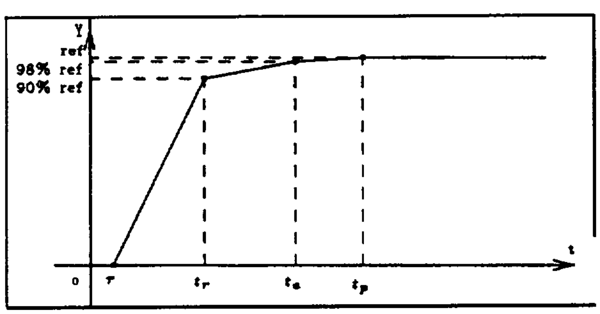


Fig. 6. Rough reference model

Results obtained with our two fuzzy controllers will be compared using the following criteria:

- a quadratic position error index: $J(\epsilon_r) = \int_0^t q_{11} \cdot \epsilon_r^2 . dt$
- a quadratic angle error index: $J(\epsilon_\Phi) = \int_0^t q_{22} \cdot \epsilon_\Phi^2 . dt$
- a quadratic input cost index: $J(u_S) = \int_0^t u_S^2 . dt$

A. Sensory fusion approach

A logical way of fusing the input variables is given by:

$$\begin{cases} X_\Phi = a\Phi + b\Delta\Phi \ \ (a > 0,\ b > 0) \\ X_e = ce + d\Delta e \ \ (c > 0,\ d > 0) \end{cases}$$

If X_e is zero, this means that the cart is stable in setpoint, or that it is moving towards it. The same holds for X_Φ. If X_Φ is zero the angle is stabilized or approaching zero. Thus, stabilizing X_Φ and X_e stabilizes the system. The larger the absolute value of X_Φ and X_e, the more the position of the cart and the angle are in 'danger'. Therefore, X_Φ and X_e can be referred to as the angle emergency and the position emergency, respectively [3].

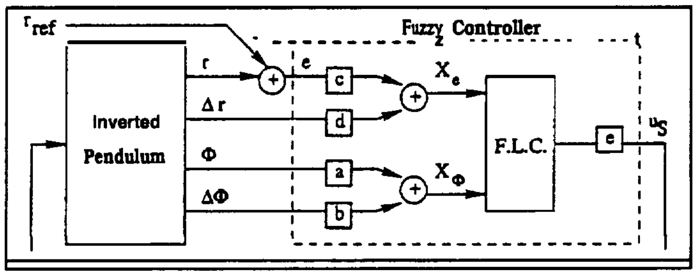


Fig. 7. FLC **based on sensory fusion**

One can easily write the following five rules to control the pendulum both in position and angle:

R^1: If X_Φ is Negative Then *us is* Negative.
R^2: If X_Φ is Positive Then u_S is Positive.
R^3: *If* X_Φ *is* Zero And X_e is Negative Then u_S is Negative.
R^4: If X_Φ is Zero And X_e is Zero Then u_S is Zero.
R^5: If X_Φ is Zero And X_e is Positive Then us is Positive.

Let us examine the third rule (R^3) in more details: *'If* X_Φ *is Zero And* X_e is *Negative Then us* is *Negative"*. This negative control action implies X_Φ >0 and therefore (Rule R^2) *us* becomes positive in order to balance the pole. The position of the cart increases as desired.

The five previous rules can be written in a more compact form in the following rule base table:

X_Φ		N	Z	P
X_e	N	N	N	P
	Z		Z	
	P		P	

TABLE I

FLC RULE-BASE $\rightarrow u_S$

The evolution of the input/output gains and evolution of the performance criteria are given in the following table:

FLC based on sensory fusion	Input/Output gains					Performance criteria		
	a	b	c	d	e	$J(\epsilon_r).10^{-2}$	$J(\epsilon_\Phi).10^{-2}$	$J(w)$
Initialized	11.45	2	1	2	5	7.32	6.82	1.37
Tuned	11,98	2.33	0.71	1.95	5.04	0.93	3.18	0.7948

TABLE II

EVOLUTION OF *FLC* INPUT/OUTPUT GAINS AND THE PERFORMANCE CRITERIA

The response to a 30 cm position reference before and after tuning is given in Fig. 8. Performances obtained with the five rules FLC are very satisfying.

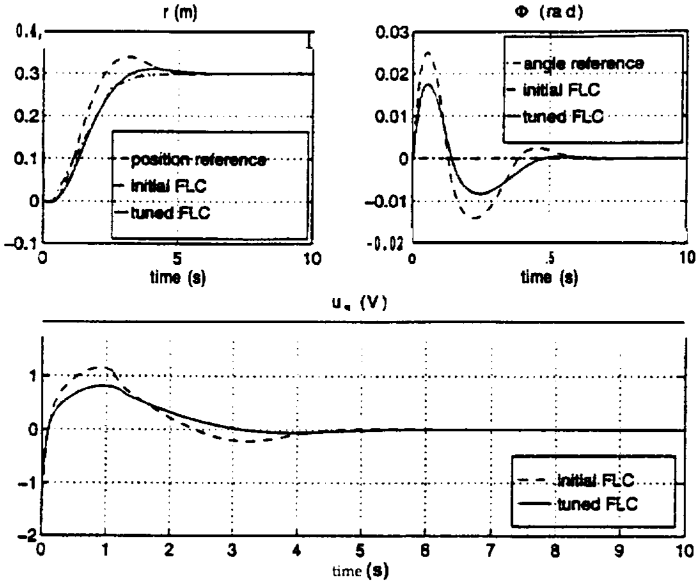


Fig. 8. FLC **based on sensory fusion**

B. Hierarchical structure

The chosen hierarchical structure (see Fig. 9) introduces a second FLC, resulting into two FLCs connected in series.

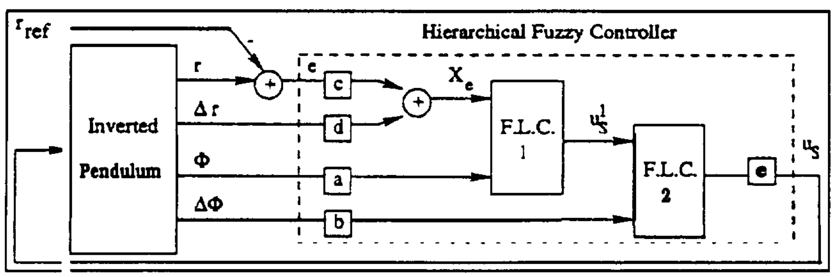


Fig. 9. Hierarchical fuzzy controller

FLC_1 combines the position emergency X_e and the angle Φ to calculate a first control signal. FLC_1 rule base is similar to the one written for the sensory fusion based FLC. The function of FLC_2 is to keep the pole balanced by calculating the correct control signal associated with the previous control value and the angle velocity. FLC_1 rule base and FLC_2 rule base are given in Table III.

Φ	N	Z	P
X_e N	N	N	P
X_e Z	N	Z	P
X_e P	N	P	P

FLC_1 **rule-base** $\rightarrow u_S^1$

$\Delta\Phi$	N	Z	P
u_S^1 N	N	N	Z
u_S^1 Z	N	Z	P
u_S^1 P	Z	P	P

FLC_2 rule-base $\rightarrow u_S$

TABLE III

HIERARCHICAL FLC RULE-BASES

The evolution of the input/output gains and evolution of the performance criteria are given in Table IV.

Hierarchical FLC	Input/Output gains					Performance criteria		
	a	b	c	d	e	$J(\epsilon_r).10^{-2}$	$J(\epsilon_\Phi).10^{-2}$	$J(u_S)$
Initialized	11.45	2	1	2	5	8.81	7.51	1.4877
Tuned	12.09	2.24	0.68	1.91	5.05	0.79	3.08	0.786

TABLE IV

EVOLUTION OF *FLC* INPUT/OUTPUT GAINS AND THE PERFORMANCE CRITERIA

The response to a 30 cm position reference before and after tuning with the hierarchical fuzzy controller (14 rules) is given in Fig. 10. One *can* notice that this response is very similar to the one obtained with the FLC based on sensory fusion. The results are very slightly improved compared to the previous FLC containing only five rules (compare Table II and Table IV).

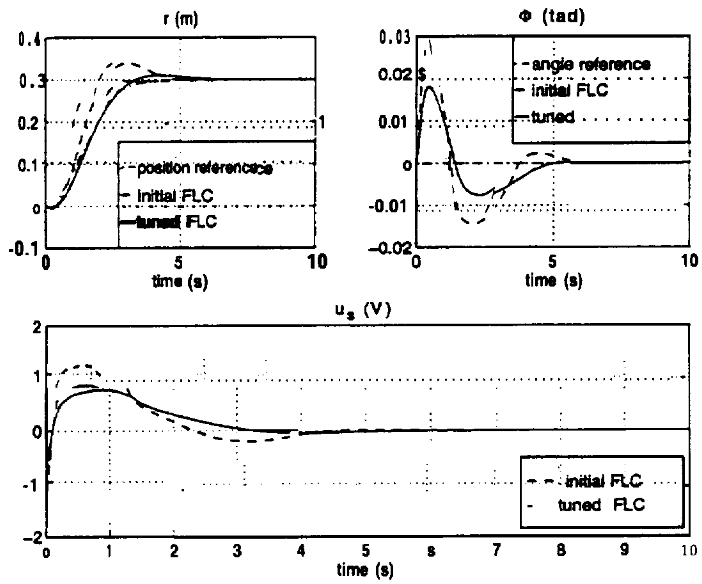


Fig. 10. Hierarchical FLC

IV. Conclusion

In fuzzy control, using reduction methods and learning techniques can help the control designers. When facing large scale system it is interesting to reduce the FLC complexity in order to facilitate its synthesis using fusion and hierarchy. Satisfactory results have been obtained on the inverted pendulum with only five rules.

References

[1] M. Jamshidi. **Large-scale** *systems modelling, Control and Fuzzy Logic.* **Prentice-Hall, Englewood Cliffs, U.S.A., 1996.**
[2] G.V.S. Raju and Jun Zhou. **Fuzzy logic process controller. In** *IEEE International Conference on Systems Engineering,* **pages 145-147, Pittsburgh, USA, August 1990.**
[3] T. Yamakawa. *Wine-glass balancing problem.* Kyushu **Institute of Technology,** Iizuka, Fukuoka, Japan, **1991. Videotape.**

URC97080

Validity using pump-probe pulses to determine the optical response of niobate crystals

Huimin Liu and Weiyi Jia

S80/76

Department of Physics, University of Puerto Rico at Mayaguez
Mayaguez, PR 00681, U.S.A.

1. INTRODUCTION

A variety of niobate crystals have found their places in nonlinear optical applications as well as in laser devices. In recent years much attention has been paid to study the ultrafast optical response in a variety of photorefractive crystals such as $KTa_{1-x}Nb_xO_3$ and $KNbO_3$ crystals[1,2], glasses[3], semiconductors and polymers[5] for applications in optical switching, information processing, optical computing, and all-optical device systems. Third-order optical nonlinearity is the most important property for realization of all-optical switching. Therefore experiments have been performed on the third order susceptibility using a variety of techniques such as the third-order harmonic generation, EFISH and degenerate four-wave mixing(DFWM). The latter has been conducted with a variety of pump wavelengths and with nanosecond, picosecond and femtosecond pulses.

Niobate crystals, such as potassium niobate $KNbO_3$, potassium tantalate niobate KTN family ($KTa_{1-x}Nb_xO_3$), strontium barium niobate SBN ($Sr_xBa_{1-x}Nb_2O_6$) and potassium-sodium niobate SBN (KNSBN) are attractive due to their photorefractive properties for application in optical storage and processing. The pulsed probe experiments performed on theses materials have suggested two types of time responses. These responses have been associated with an coherent response clue to $\chi^{(3)}$, and a long lived component due to excited state population. Recent study of DFWM on $KNbO_3$ and KTN family reveals that the long lived component of those crystals depends on the crystal orientation, A slowly decaying signal is observable when the grating vector $\mathbf{K_g}$ is not perpendicular to the C-axis of those photorefractive crystals', otherwise the optical response signal would be only a narrow coherent peak with FWHM equal to the cross-correlation width of the write beam pulses. Based on this understanding, we study the photodynamical process of a variety of niobate crystals using DFWM in a $\mathbf{K_g} \perp C$ geometry with a ps-YAG:Nd laser operating at 532nm. However, the discrepancies in numerical estimations of $\chi^{(3)}$ in these materials and other nonlinear optical media[6] have resulted in a number of discussions[7,8]. In order to better understand the photodynamical process of niobate crystals after an short pulse excitation we analyze the factors governing the coherent signal and present the DFWM spectra of niobate crystals.

2. EXPERIMENTAL

Host materials SBN:61, KNaSBN:75 (potassium-sodium SBN) and mixed KTN ($KTa_{1-x}Nb_xO_3$) niobate crystals with $O \leq x \leq 0.84$ and $x=1$ were used in the experiment. The samples were poled along the C-axis, and cut with the C-axis along one of the surfaces. A single 25 ps pulse from a mode-locked, Q-switched Nd:YAG laser operating at 10 Hz was frequency doubled to $\lambda=532$ nm and split into three pulses. These pulses were then spatially overlapped in the sample in the conventional backward propagating degenerate four-wave mixing geometry (see Fig. 1) with the counter propagating beam acting as the probe. The pulses were co-linearly polarized, with the pump beam/probe beam intensity ratio of 10:1. The scattered light was monitored by a Molectron Joule meter and the output was stored in a computer. The arrival of the probe beam relative to the simultaneously arriving pump beams was controlled by a computer driven optical delay line, At the sample, the three beams had a nearly Gaussian spatial profile with a radius of 250 μm for the pump beams and 200 μm for the probe beam. The smaller probe beam radius reduces effects due to the spatial nature of the induced grating, particularly, effects arising from the Gaussian profile of the pump beams and variations due to pump beams crossing angle 2θ. The temporal and spatial overlap of the three beams at zero delay was determined by optimization

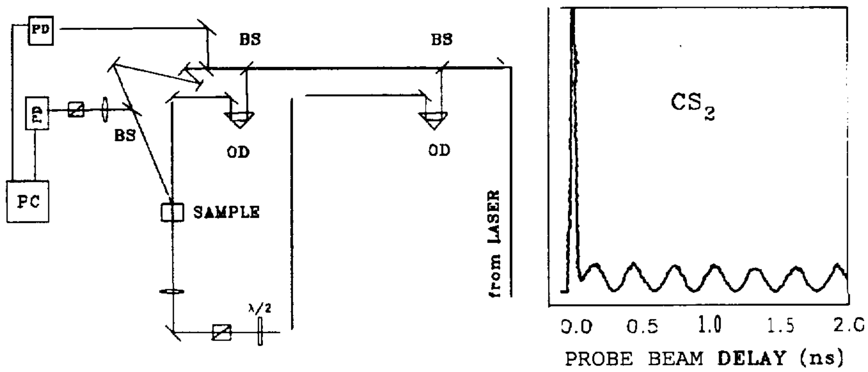


Fig. 1. Setup of DFWM experiment using CS_2 as a standard,

of the instantaneous response signal from a CS_2 sample. The value of 2θ was varied between 3.20-30.4°. The polarization configuration of the four pulses including signal was (ssss). Further details of the laser system and optical delay lines are given in refs[1].

3. DFWM spectra measurement

3.1 DFWM spectra of host niobate crystals

Fig. 2 is the time-resolved DFWM spectrum obtained from $KNbO_3$ crystal at the different orientation angle ϕ, where ϕ is the angle between K_g and C-axis of the crystal. The spectrum of +=900 also represents a common feature of instantaneous optical response signal of all poled (single domain) SBN and KSBN niobate crystals in the $K_g \perp C$ geometry. This signal is associated with the third-order susceptibility of the material. It arises from fast, coherent processes. The FWHM of the signal is close to the autocorrelation width of three pulses, similar to that of a CS_2 reference sample. No signal broadening is observed. The probe pulse, with intensity equal to 10% of the total laser intensity, was passed through an optical delay line, and was incident on the sample in a backward configuration, The sophisticated optical delay line used can reach an accuracy of 15 fs. When crystal is rotated or ϕ decreases a slowl y growing response signal observed starting from $\phi = 75°$. The spectrum is now composed of two signals: a sharp coherent signal with a peak at zero-delay, and a slowly rising signal component. The latter reaches its maximum intensity at $\phi = 0°$ ($K_g // C$). It reflects movement of the dynamic lattice due to creation of free-carriers, In the bright region of the grating produced by crossing two write-pulses inside the sample the carriers are created in the conduction band which consists of $d\xi$ orbitals of Nb electrons, leaving behind holes in the valence band, which consists of $p\pi$ orbitals of oxygen electrons. Under the action of intense internal electric field (single domain crystal) induced in the poling process along the C-axis, those free-carriers are caused to drift along the C-axis direction until trapped at Nb^{5+} sites, forming Nb^{4+} or so-called Nb^{4+}-hole small polarons[9]. They cause a temporal lattice distortion in C direction, and consequently alter the index of refraction along K_g by a factor of $(\cos\phi)$. As ϕ deceases to O", ($K_g // C$ geometry), the coherent signal component is superimposed with the long lived slow component resulting in signal width completely masked. On the other hand, if the geometry is $K_g \perp C$ or $\phi = 90°$, however, the resultant lattice distortion has no contribution along the K_g direction, and thus no slow component could be observed.

The change in susceptibility $\Delta\chi$ is related to An, and the change in the index of refraction is:

$$\Delta\chi = \Delta[(n + i\ kc/2\omega)^2 - 1] \tag{1}$$

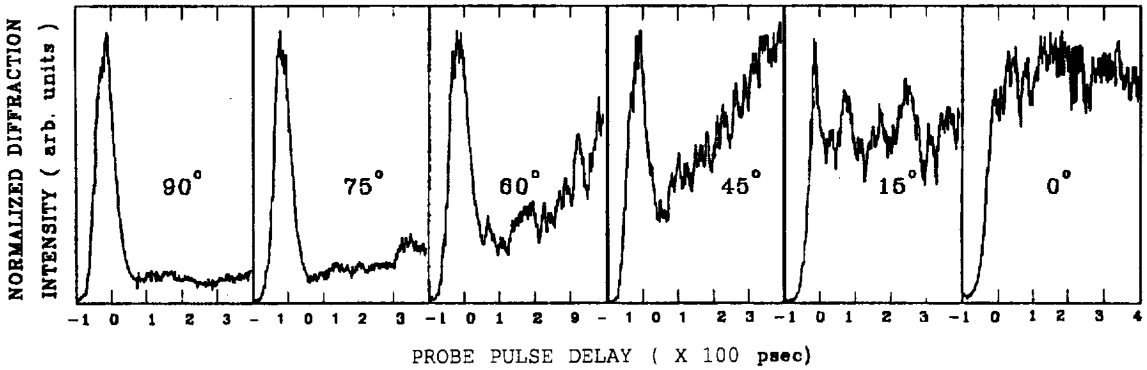


Figure 2. Time-resolved **DFWM** spectra of poled **$KNbO_3$** crystal, A narrow but intense signal at zero-delay of probe pulse represents the coherent optical response.

where ω is the angular frequency of the write-beams. $\Delta\chi$ is also a function of lattice distortion **Δr**, and is expressed as

$$\Delta\chi = \Delta N(3r^2/\xi + 1)\Delta r \cos\phi. \qquad (2)$$

N is the population density, r, the lattice constant of the crystal. ξ is a crystal dependent parameter, **Δr**, in principle, can be determined from luminescence measurements. Expression (2) reflects the crystal orientation dependence of the observed **DFWM signal** intensity. **In Fig.2(c),** where **K_g//*C* (ϕ=0),** the change in susceptibility due to temporal lattice distortion reaches a maximum. This part of the contribution to susceptibility **change** is added **to** that due **to** the instantaneous grating. **Since** the space charge **field** is parallel to the C-axis and the lattice distortion occurred also along the C-axis, no time delay is needed for those carriers making a realignment, unlike what we have seen in figure l(b).

3.2. ADJUSTMENT OF PUMP-PULSES **IN DFWM**

Under a good experimental condition the obtained **DFWM** spectrum of the CS_2 standard shows a sharp coherent peak **at** zero time delay followed by an oscillating **signal as** shown in the insert of Figure 1. The relative delay between the two write pulses was carefully adjusted by monitoring the peak intensity of CS_2 reference sample. A subtle variation in the relative delay was found to give a **significant** change in the obtained **DFWM** spectra. As shown in Fig. 2, the coherent signal of CS_2 changes its profile as well as the peak intensity if the relative delay between the two pump pulses changes. The peak intensity of the signal at the zero delay is -15 times stronger than that at **-60ps** (pulse B arrives in advance) or at -I- **60ps** (pulse A arrives in advance). Its width **(FWHM)** is about a **half** of the width at **60ps** delay. The observed dependence of the **DFWM** signal on the relative delay times between the two write beam pulses cart be described as follows. The two write-pulses are assumed to have Gaussian temporal distributions with **pulse** widths $2\tau_p$ and to cross spatially inside the sample. The intensity of pulse A is expressed as

$$I_A = I_A^0 \exp[-(\frac{t}{\tau_p})^2] \qquad (3)$$

where t =0 is defined at the pulse peak. if the time delay between the two write-pulses is At, the intensity of pulse B at time t is therefore written as

$$I_B = I_B^0 \exp [- (\frac{t+ \Delta t/2}{\tau_p})^2] \qquad (4)$$

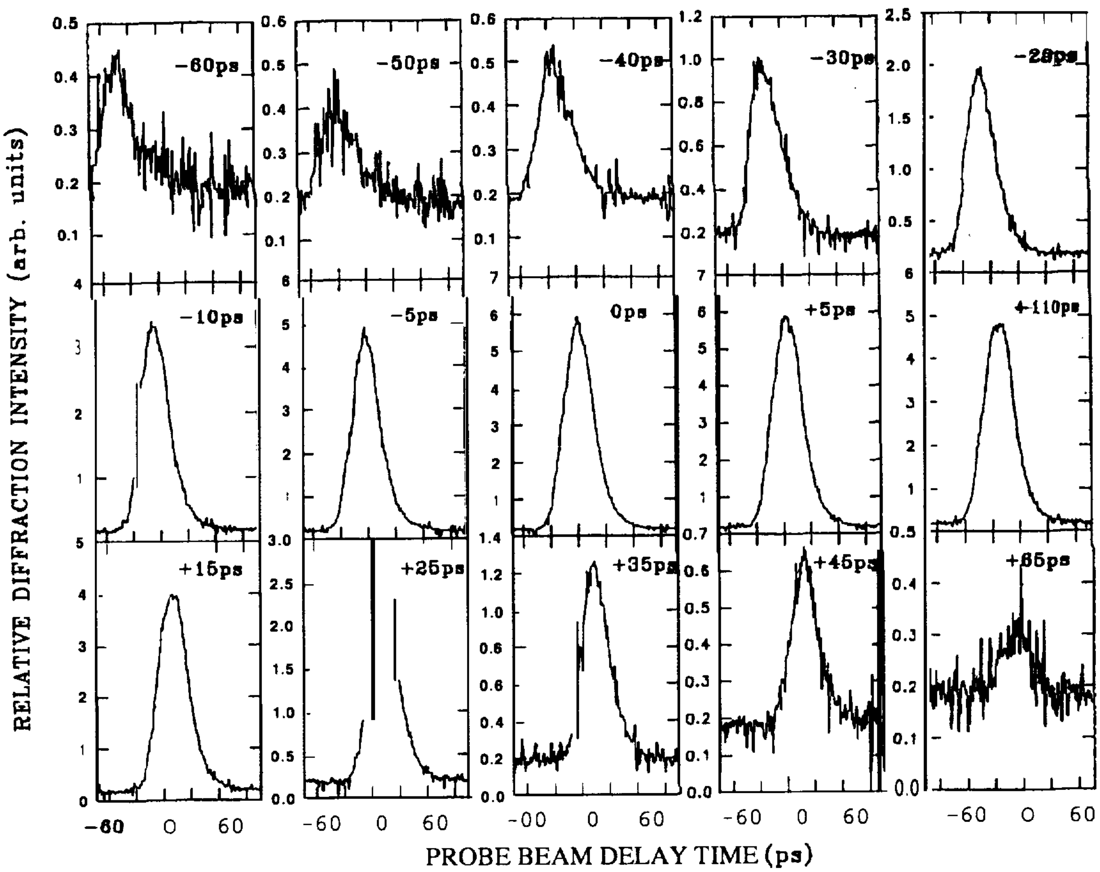


Fig. 3. Diffraction signal intensity as a function of the relative delay between the two pump pukes for CS_2 standard.

The intensity at the peak of the interference pattern between the two write-pulses can be expressed as:

$$I = I_A(t) -1- I_B(t) + AI \qquad (5)$$

where, $$\Delta I = AA\ A_B \exp[-(\frac{\Delta t}{2\tau_p})^2] \exp[-(\frac{t}{\tau_p})^2] \qquad (6)$$

and AA , A_B are the magnitudes of the electric fields of the pump beams A and B respectively. Taking $I_A = I_B$, the interference intensity is then expressed as

$$I = I_A \{1 + \exp[-(\frac{\Delta t}{2\tau_p})^2] + 2\exp[-3/2(\frac{\Delta t}{2\tau_p})^2]\} \qquad (7)$$

Using the power-law dependence measured for the diffraction efficiencies of CS_2,[10] the normalized diffraction intensities versus At are found to be in good agreement with Fig.3. Eq.(7) shows how important the fine adjustment of At is. If the laser pulse width used in DFWM is in nanosecond range the misalignment At of the two write pulses arising from the path difference is not significant. If the pulse width is in picosecond or femtosecond range, however, a little inadvertency may cause a significant discrepancy in χ^3 measurement.

4. CONCLUSION.

In DFWM measurement of niobate crystals the host excitation caused by two-photon absorption may form small polarons in samples, resulting in a broader, long lived, and slowly decaying signal. Using DFWM in a K_g 1 C geometry the signal broadening of the coherent response can be eliminated. Regarding to $\chi^{(3)}$ measurement for different niobate crystals, the pump-pulse adjustment in DFWM measurement is crucial. It can he optimized using a CS_2 reference sample.

Acknowledgements
This work is supported in part by NASA under the contract NASA-MURC NCCW-0088 and the U.S. Army Research Office, DAAH04-96-10416.

References

1) H. Liu, R.C. Powell, and L.A. Boatner, Phys.Rev.B, 44 (1991) 2461; 49 (1994) 6323; Opt. Mater., 4 (1995) 691.
2) M. Zgonik and P. Gunter, J. Opt. Soc.Am.B, 13 (1996) 570.
3) N. Sugimoto, H. Kanbara, S. Fujiwara and K. Tanaka, Opt. Lett., 21 (1996) 1637.
4) A.E. Bieber, D.F. Prelewitz, T.G. Brown and R.C. Tiberio, J. Opt. Soc.Am.B, 13 (1996) 34.
5) V.L. Bogdanov and A.G. Spiro, SPIE Proc., 2527 (1995) 275.
6) H.Liu, B. Taheri, W.Jia, Phys.Rev.B, 49 (1994) 10166.
7) W. J. Blau, H. J. Byrne, D. J. Cardin, T. J. Dennis, J. P. Hare, H. W. Kroto, R. Taylor and D. R. M. Walton, Phys.Rev.Let., 67 (1991) 1423.
8) Z. H. Kafafi, J. R. Lindle, R. G. S. Pong, F.J.Bartoli, L. J. Lingg and J. Milliken, Chem.Phys.Let., 188 (1992) 492.
9) O.F. Schirmer and D. von der Linde, Appl. Phys. Lett. 33 (1978) 35.
10) M.G. Kuzyk, R.A. Norwood, J.W. Wu, and A.F. Garito, J. Opt. Soc.Am.B, 6 (1989) 154.

Page intentionally left blank

URC97081

A Categorization of Dynamic Analyzers

Michelle R. Lujan*
Department of Computer Science
The University of Texas at El Paso
El Paso, TX 79968
email mlujan@cs.utep.edu

1 Introduction

Program analysis techniques and tools are essential to the development process because of the support they provide in detecting errors and deficiencies at different phases of development. The types of information rendered through analysis includes the following: statistical measurements of code, type checks, dataflow analysis, consistency checks, test data verification of code, and debugging information. Analyzers can be broken into two major categories: dynamic and static. Static analyzers examine programs with respect to syntax errors and structural properties [17], This includes gathering statistical information on program content, such as the number of lines of executable code, source lines. and cyclomatic complexity. In addition, static analyzers provide the ability to check for the consistency of programs with respect to variables. Dynamic analyzers in contrast are dependent on input and the execution of a program providing the ability to find errors that cannot be detected through the use of static analysis alone. Dynamic analysis provides information on the behavior of a program rather than on the syntax [1 7]. Both types of analysis detect errors in a program, but dynamic analyzers accomplish this through run-time behavior.

This paper focuses on the following broad classification of dynamic analyzers:

- Metrics
- Models
- Monitors

Metrics are those analyzers that provide measurement. The next category, models, captures those analyzers that present the state of the program to the user at specified points in time. The last category, monitors, checks specified code based on some criteria. The paper discusses each classification and the techniques that are included under them. in addition, the role of each technique in the software life cycle is discussed. Familiarization with the tools that measure, model and monitor programs provides a framework for understanding the program's dynamic behavior from different, perspectives through analysis of the input/output data.

2 Metrics

The term *metrics* is defined as the measure of properties of systems [14]. Metrics are used to measure the quality of a program, control productivity of software projects, and attempt, to predict the effort in developing software [2]. In addition, errors can be detected in programs through the use of statistical information. According to [14], metrics should have the following properties to be useful:

1. Automatable: the metrics must be processed automatically.
2. Feasible: the cost of gathering metrics must be feasible in cost and time in order to be useful.
3. Understandable: the metrics must have theoretical and empirical foundations and clear meaning.

* This work was sponsored by NASA under contract NAG-1012 and NCCW-0089

4. Sensitive: the metrics should be sensitive to all the factors that affect the quality to be estimated and not those that, are unrelated.

5. Applicable: the metrics should be applicable to any programming language and to a large class of systems, and not to any particular stage of the lifecycle.

6. Useful: metrics should not measure values but should provide feedback concerning software activities.

7. Flexible: a variety of metrics is needed to provide users with a broad picture of the properties of the system.

W bile static metrics generate information about variable anomalies and program content, dynamic metrics produce data on program interconnections and variable behaviors [9]. one category of dynamic metrics, coverage analyzers, records information concerning logic, statements, cent rol paths, decision, conditions, dataflow and iteration. A few are discussed next. ***Control path coverage analyzers*** record the control paths covered by each test case [5]. ***Statement coverage analyzers*** check that every stat ement in the program is executed by the test data set at least once. ***Decision coverage analyzers*** verifies that each predicate decision assumes a ***true*** and a ***false*** outcome at least once during testing on a test suite. Clearly, coverage measures are categorized as dynamic analyzers because they require execution of the program in order to collect the metrics. Even though the general notion of coverage provides a simple methodology for developing test suites, many errors may still escape detection [17].

Reliability measures, another category of dynamic metrics, generate a probability measure that a software fault does not occur during a specified time interval [1 4]. These measures provide information concerning the number of failures during a specified amount of runs and the totality of failures within a specified interval.

Performance measures produce a statistic on the overall performance of a program. Measures of this type include estimating the stability and reliability of a program [9] [14]. An example of a performance measure that approximates the stability of a program is Yau and Collofello's Logical Stability Metric [9]. This calculation is used to determine the expected impact of a modification of a variable in a module. The stability metric reflects the result of changing a single variable in a module and the impact of the change on the behavior of the program. The authors' basic argument is that maintenance breaks down to changes in variables regardless of how complex the task.

3 Models

Observing the state of a program throughout different points of execution time provides valuable information to the person analyzing a program. ***Models*** provide a systematic method to accomplish this end. By supplying the users with the ability to test, hypothesis and draw conclusions about program behavior, models ***support*** the evolution of programs over time [3].

Interpreters, a category of models, execute and translate programs at the same time. This is an advantage because many interpreters show the likely place and cause of errors [13]. Interpreters are made up of four basic components [16]:

1. an engine to interpret the program,

2. memory that contains the pseudocode to be interpreted,

3. representation of the control state of the interpretation engine, and

4. a representation of the current state of the program being simulated.

As a result, the state of the machine can be observed through the execution of the program with the provided input. Because interpreters depend on execution and input, they are categorized as dynamic analyzers.

Prototypes, the second category of models, provide a mechanism to learn more about the problem and the problem solution through a partial implementation of the system [4]. Prototypes enable potential users to experiment with the system relatively early in the development process. Essent ially, this allows users to provide feedback to designers on whether the behavior of the system is as expected and to determine user needs [4]. Prototyping is often viewed as a way of progressively developing an application and at the same time understanding the requirements [8]. The cost, of the system is reduced because prototypes can be produced early in the development process without implementing the entire system. Prototypes are viewed as the first version of the system. Two types of prototypes are throwaway and evolutionary. Throwaway

prototypes are those that are discarded after it is used. Evolutionary prototypes are those that continually changed over time until it behaves as expected (becoming the final producet) [4]. The evolutionary prototype is more cost effective because it is not destroyed after it is used [4]. A prototyping environment is provided through Computer-Aided Prototyping (CAPS) [i l], being developed at the U.S. Naval Postgraduate School. CAPS is comprised of the following tools that aid the development of a prototype: a graph data model, change merging facility, automatic generators for schedule and control code and automated retrievals for reusable components.

Debuggers aid in locating, analyzing, and correcting errors by providing the user with the ability to examine a program by executing code one line at a time. These tools allow the users to inspect the execution of a program in detail, to control the time that each instruction takes to execute, and to control the progress of the computation [3]. The history of execution can be generated along with the state of variables and the machine. This trace facilitates the collection and manipulation of information [3].

A few examples of debuggers are YODA [12], TSL [12], and EBBA Event Based Behavioral Abstraction) [1]. The YODA system stores Ada event histories as Prolog facts. Predicates in Prolog define the common temporal relationships. This system has the capability to specify when variables can be updated, what, values variables can take on when updated and the communication between variables. The TSL system automatically checks specifications against the events produced by an Ada tasking program. In addition, it uses Ada semantics to ensure that pairs of events appear in the correct order in the event history. These debuggers provide a method to detect errors in concurrent Ada programs [1 2].

The debugging tool EBBA is also based on event histories. A distinguishing feature of this approach is its ability to model system behavior through clustering and filtering. Clustering expresses behavior as composite events whereas filtering removes from consideration those events that are not needed for the behaviors being investigated [1].

4 Monitors

Programs often need to meet certain criteria in order to provide the desired functionality. *Monitors* provide the ability to ***examine code*** against criteria imposed by the user or the designer to check for satisfiability. The aim is to monitor the reliability and quality of software systems.

The first, category of monitors, assertion checkers, are those tools that support automatic runtime detecion of software faults during debugging, testing and maintenance [15]. Through the use of assertion checkers, developers are provided with the capability to incorporate assertions in programs in order to ensure that they are not violated throughout execution. Assertions are defined as formal specifications that describe the properties of programs using mathematical notation. In other words, assertions specify what a system is supposed to do instead of how to implement it. Assertion checkers verify that assertions are maintained throughout, runtime. A few tools that fall under the category of assertion checkers follow.

The annotation language ANNA (A NNotated Ada) is used to embed assertions into Ada programs and performs consistency tests to determine if the computation satisfies the specified properties, ANNA has the ability to ensure that assertions are maintained throughout the execution of a program. Features of A NNA include the following [10]:

- generates consistency checks from annotations on types, variables, subprograms, and exceptions,
- uses incremental theorem proving to check algebraic specifications at runtime, and
- constructs large software systems based on algebraic specification of system models.

Based on ANNA, the Annotation Preprocessor (APP) for C programs is a replacement for the standard preprocessing pass of the C compiler. In addition, APP provides a mechanism to define how assertion violations will be handled during execution and the level of checking that is to be done [15]. An assertion in APP specifies a constraint that is related to some state of computation. Constraints are specified using C's expression language. APP converts each assertion into a runtime check in order to test for violations of constraints. In this way, APP provides a convenient method to specify and maintain assertions.

FORMAN (FORmal Annotation), an assertion language, has the capability to express assertions on events and sequences of operations and events [6]. Included in FORMAN, is the ability to describe universal assertions on the program. Assertions can be collected into ii braries to increase the level of automation to encounter errors. FORMAN includes a flexible language for trace specification based on event patterns

and regular expression [1]. In addition, FORMAN has the capability to express both general operational assertions and declarative assertions.

Another language, Behavioral Expressions (BE), provides the capability to write assertions about sequences of process interactions. It also has the functionality to describe allowed sequences of events as well as some predicates [1]. Events are used to describe process communication, termination, connection, and detachment of process to channels. BE performs evaluations of assertions at runtime.

Context monitoring [7] is an approach that provides the developer with tools to manage, and communicate across personnel, application domain knowledge about the properties on and relationships between objects being modeled by a software system. Knowledge about the data, the intended context in which programs will run and other knowledge about the program is captured through integrity constraints. The constraints are elicited from domain experts, customers, analysts, designers and programmers using established methods. The constraint satisfiability mechanism dynamically monitors a program to ensure that the constraints are being enforced by the program. If a violation occurs, the user is notified and, because links exist between the constraints and the documents that support the constraint, the user can identify the source of the constraint,. This approach is distinguished from the others because the constraints are not embedded in the program code, but are maintained in a repository.

5 Summary

Classification	*Techniques*	*Principal Life Cycle Support*
Metrics	Coverage measures	Testing
	Reliability measures	Implementation Maintenance
	Performance measures	Implementation Maintenance
Models	Interpreters	Implementation Testing
	Prototypes	Requirements Design
	Debuggers	Implementation Maintenance
Monitors	Assertion checkers	Requirements Implementation/Maintenance Testing
	Context monitoring	Requirements Design Implementation/Maintenance Testing

Figure 1: A Classification of Dynamic Analyzers.

Examining runtime behavior is an important step in error detection and analysis. Data gathered from runtime behavior can provide insight into errors which may not be detected through static analysis. Metrics, models and monitors all produce different types of dynamic information about programs and, depending cm the technique, support different aspects of the software life cycle (see Fig. 1). Metrics generate statistical information about variables and program interconnections. Models monitor the state of the machine at specified times during program execution. Monitors oversee that criteria specified by designers or users are not violated, Even though dynamic analyzers alone do not supply enough data about programs to localize all errors, they do furnish information that static analyzers do not.

References

[1] Auguston, M., "A Language for Debugging Automation", in *Proceedings of SEKE*, U.S.A.: Knowledge Systems Institute, 1994, pp.108-115.

[2] Basili, V. R., Selby, R. W., Yun, T., " Metric Analysis and Data Validation Across Fortran Projects", *IEEE Transactions Software Eng.* SE-9 (6), 6.52-663 (1983).

[3] Brindle, A. F., Taylor, R. N., Martin, D. F., "A Debugger for Ada Tasking". *IEEE Transactions Software Eng.* SE-15 (3), 293-304 (1989).

[4] Davis, A. M., "Soft ware Prototyping", in Yovits, M. C., Zelkowitz, M. V. (eds.), *Advances in Computers Vol. 40.* San Diego: Academic Press, 1995, pp. 39-63.

[5] Fairley, R., *Software Engineering Concepts.* New York: McGraw-Hill Publishing Company) 1985

[6] Fritzon, P., Auguston, M., Shahmehri, N., " Using Assertions in Declarative and Operational Models for Automated Debugging", *J. Systems Software 25,* 223-239 (1994).

[7] Gates, A. Q., F. G. Fernandez and L. Romo, "Building Systems with integrity Constraints," to appear in *The Proceedings of the Second World Conference on Integrated Design and Process Technology,* December 1-4, 1996, Austin, Texas.

[8] Ghezzi, C., J azayeri, M., Mandrioli, D., *Fundamentals of Software Engineering.* New Jersey: Prentice Hall, 1991.

[9] Kafura, D., Reddy, G. R., " The Use of Software Complexity Metrics in Software Maintenance", *IEEE Transactions Software Eng* SE-13 (3), 335-343 (1987).

[10] Luckham, D. and Von Henke, F. W., "An Overview of Anna: A Specification Language for Ada," *IEEE Software,* 20(2):9-23, 1985.

[11] Luqi, Goguen, J. and V. Berzins, "Formal Support for Software Evolution," *1994 Monterey Workshop Increasing the Practical Impact of Formal Methods for Computer-Aided Software Development: Soft ware Evolution.* Monterey, CA: U.S. Naval Postgraduate School, Sept. 7-9, 1994, pp. 10-21.

[12] McDowell, C. E., Helmbold, D. P., " Debugging Concurrent, Programs", *ACM Computing Surveys* 21 (4), 593-622 (1989).

[13] Pfaffenberger, B. *Que's Computer User's Dictionary.* 4th ed. U. S. A.: Que, 1993

[14] Ramamoorthy, C. V., Prakash, A., Garg, V., Yamura, T., Bhide, A., " Issues in the Development, of Large, Distributed, and Reliable Software", in Yovits, M. C. (ed.), *Advances in Computer Vol.* .26. San Diego: Academic Press, 1987, pp. 393-443.

[15] Rosenblum, D. S., " A Practical Approach to Programming With Assertions", *IEEE Transactions Software Eng.* 21 (1), 19-31 (1995).

[16] Shaw, M., Garlan, D., *Software Architecture: Perspectives on an Emerging Discipline.* New Jersey: Prentice Hall, 1996.

[17] White, L. J., "Software Testing and Verification", in Yovits, M. C. (ed.), *Advances in Computer Vol.* 26. San Diego: Academic Press, 1987, pp. 335-391.

Page intentionally left blank

URC97082

-32/64

A Reduction in Conservatism for Convex Linear-Quadratic Simultaneous Performance Design

Robert A. Luke, Peter Dorato, and Chaouki T. Abdallah
Department of Electrical and Computer Engineering
University of New Mexico
Albuquerque, NM 87131-1358 U.S.A.
luke@eece.unm.edu
http: //www.eece.unm. edu/students/luke/

Abstract

In this paper a fixed state feedback control law which minimizes upper bounds on linear-quadratic performance measures for m distinct plants is studied. Previous work [8] by the authors demonstrated a convex semidefinite programming solution thereby guaranteeing global optimality. The present work extends that result by proposing an algorithm which reduces the conservatism of the minimum guaranteed-cost upper bounds for each of the m performance measures.

Keywords: simultaneous stabilization, simultaneous control, semidefinit e programming, state feedback control, linear matrix inequalities (LMIs)

1 Introduction

The problem considered here is the design of a fixed state feedback control law u(t) = $-Kx(t)$ which minimizes an upper bound on the performance measures

$$\mathbb{E}\left\{\int_0^\infty \left[x_j^T(t)\ Q_j x_j(t) + u_j^T(t) R_j u_j(t)\right] dt\right\} \tag{1}$$

for $Q_j > 0$ and $R_j > 0$, each associated with one of the plants described by state space equations

$$\dot{x}_j\ (t) = A_j x_j(t) + B_j u_j\ (t) \tag{2}$$

for all $j \in I_m \triangleq \{1, \dots, m\}$. In (1) the expectation operator $\mathbb{E}\{\cdot\}$ is taken over random initial conditions satisfying $\mathbb{E}\{x(0)\} = 0$ and $\mathbb{E}\{x(0)x^T(0)\} = I$. We refer to this as a simultaneous *performance design problem.*

The paper begins with the convex reformulation of the Chang and Peng [3] guaranteed-cost control met hod discussed in Luke, et al. [8]. The problem can be solved, albeit conservatively, using widely available semidefinite programming software (El Ghaoui, et al. [5]; Nesterov and Nemirovskii [9]; and Vandenberghe and Boyd [11]). Further, the property of convexity guarantees that any solution found will be *globally* optimal. However the guaranteed-cost upper bounds on performance measures (1) were kept invariant over all systems $j \in I_m$ in order to achieve that convexity. These invariant bounds necessarily introduce conservatism to the solution and the goal of the present work is to reduce conservatism while retaining problem convexity.

The problem appears to be approachable in at least two ways. The first involves a search for a parameterization of the simultaneously stabilizing gain matrix K such that nonconvexity does not arise. This effort would be similar to that suggested by Kar [7] in a somewhat different context. But it may limit the reduction in conservatism because the search would be constrained to the set of all variable gain matrices K satisfying the pattern imposed by the parameterization. The second method involves a decomposition of the larger nonconvex problem into a set of two convex, but coupled, systems. We deal with the latter option and follow a direction suggested by the work of Grigoriadis and Skelton [6]. This course of action also

tends to limit the reduction in conservatism and this limiting action is analyzed herein. Finally an example frequently appearing in the literature is used to demonstrate a significant reduction in the guaranteed-cost upper bounds, compared to what was previously achieved.

2 Convex Guaranteed-Cost Control

Using the methods of guaranteed-cost control developed by Chang and Peng [3], it is possible to show (Dorato, et al. [4]) that an upper bound on the performance measures (1) involving the transient state response $x_j(t)$ and the control effort $u_j\ (t)$ are optimized for each of the systems by assigning to each an integral quadratic performance function (1). The expectation operator taken over all initial conditions z(O) results in

$$\mathbb{E}\left\{\int_0^\infty \left[x_j^T(t)Q_jx_j(t) + u_j^T(t)R_ju_j(t)\right] dt\right\} \leqslant \mathrm{tr}\ \{P\}\ ,\forall j \bullet I_m \tag{3}$$

where the single matrix $P = P^T > 0$ satisfies each of the m matrix Lyapunov inequalities

$$(A_j - B_jK)^T P + P\,(A_j - B_jK) + Q_j + K^T R_j K < 0 \tag{4}$$

for all $j \in I_m$. The existence of such a $P = P^T > 0$ is also sufficient to guarantee that the state feedback control law $u_j\ (t) = -Kx_j(t)$ simultaneously stabilizes the m distinct plants (2). See Boyd, et al. [2, pp. 100-101]. It turns out that this guaranteed-cost control problem can be reduced to a convex programming problem, the advantage being that any converged solution is guaranteed to be *globally* optimal. Consider the change of variables used by Bernussou, et al. [1]: for some real matrix $Y = Y^T > 0$, let $P = Y^{-1}$ and $K = XY^{-1}$. An equivalent convex programming problem can be formulated as the following Optimization Problem 2.1 (see [8]),

Optimization Problem 2.1 (Conservative Gain)

$$\min_{X,Y,Z}\ tr\ \{Z\} \tag{5}$$

for real matrix variables $Y = Y^T > 0$ and $Z = Z^T > 0$, subject to $m + 1$ separate linear matrix inequality constraints

$$\begin{bmatrix} -YA_j^T - A_jY + B_jX + X^TB_j^T & Y & X^T \\ Y & Q_j^{-1} & 0 \\ X & 0 & R_j^{-1} \end{bmatrix} > 0\ , \quad \begin{bmatrix} Z & I \\ I & Y \end{bmatrix} > 0 \tag{6}$$

for all $j \in I_m$. The optimal gain is calculated as $K^* = X^*\,(Y^*)^{-1}$.

An example found in Paskota, et al. [10]; Luke, et al. [8], and others is used to demonstrate Optimization Problem 2.1 and to produce a point for comparison with results of reduced conservatism to follow in Section 3. A static state feedback gain K is sought which simultaneously stabilizes four different operating points of an airplane trajectory in the vertical plane. The operating points are specified with a set of m = 4 state differential equations (2) assuming a scalar input u. The state coefficient matrices are

$$A_1 = \begin{bmatrix} -0.9896 & 17.41 & 96.15 \\ 0.2648 & -0.8512 & -11.89 \\ 0 & 0 & -30 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -0.6607 & 18.11 & 84.34 \\ 0.08201 & -0.6587 & -10.81 \\ 0 & 0 & -30 \end{bmatrix}, \tag{7}$$

$$A_3 = \begin{bmatrix} -1.702 & 50.72 & 263.5 \\ 0.2201 & -1.418 & -31.99 \\ 0 & 0 & -30 \end{bmatrix}, \quad A_4 = \begin{bmatrix} -0.5162 & 26,96 & 178.9 \\ -0.6896 & -1.225 & -30.38 \\ 0 & 0 & -30 \end{bmatrix}, \tag{8}$$

and the control coefficient vectors b_j *are*

$$b_1 = \begin{bmatrix} -97.78 \\ 0 \\ 30 \end{bmatrix}, \quad b_2 = \begin{bmatrix} -272.2 \\ 0 \\ 30 \end{bmatrix}, \quad b_3 = \begin{bmatrix} -85.09 \\ 0 \\ 30 \end{bmatrix}, \quad b_4 = \begin{bmatrix} -175.6 \\ 0 \\ 30 \end{bmatrix}.$$

The state coefficient matrices Q_j and control coefficient matrices R_j used in objective functions (1) are each set to an identity matrix of appropriate dimensions. Using numerical interior point programming methods discussed by Nesterov and Nemirovskii [9], LMITOOL by El Ghaoui, et al. [5], and the semidefinite programming software package SP by Vandenberghe and Boyd [11], the results turn out to be

$$X^* = \begin{bmatrix} -0.2593 \\ 0.0061 \\ 0.0560 \end{bmatrix}^T, \; Y^* = \begin{bmatrix} 3.3514 & -0.3781 & -0.1683 \\ -0.3781 & 0.0569 & 0.0208 \\ -0.1683 & 0.0208 & 0.0387 \end{bmatrix}, \; Z^* = \begin{bmatrix} 1.2380 & 7.7939 & 1.2021 \\ 7.7939 & 70.9534 & -4.1972 \\ 1.2021 & -4.1972 & 33.3622 \end{bmatrix},$$

indicating an optimal performance matrix bound of

$$\mathrm{tr}\{P^*\} = \mathrm{tr}\left\{(Y^{-1})^*\right\} = 105.3. \tag{9}$$

The optimal gain vector $K^* = X^* (Y^*)^{-1}$ is

$$K^* = [\; -0.2063 \;\; -1.8247 \;\; 1.5305 \;], \tag{10}$$

3 Distinct Performance Bounds

Now the use of distinct scalar functions $\mathrm{tr}\{P_j\}$ is addressed to bound each of the integral quadratic performance indices (3). Consider the simultaneously stabilizing gain matrix K satisfying $K = X_1 Y_1^{-1} = \ldots = X_m Y_m^{-1}$ and rewrite Optimization Problem 2.1 as the following Optimization Problem 3.1.

Optimization Problem 3.1 (Nonlinear Problem)

$$\min_{K, \{(Y_j, Z_j)\}_{j \in I_m}} \sum_{j=1}^{m} tr\{Z_j\} \tag{11}$$

for real matrix variables K *and* $\{(Y_j, Z_j) : Y_j = Y_j^T > 0, Z_j = Z_j^T > 0\}_{j \in I_m}$, *subject to* $2m$ *separate* *nonlinear* matrix *inequality constraints*

$$\begin{bmatrix} -Y_j A_j^T - A_j Y_j + B_j K Y_j + Y_j K^T B_j^T & Y_j & Y_j K^T \\ Y_j & Q_j^{-1} & 0 \\ K Y_j & 0 & R_j^{-1} \end{bmatrix} > 0 \, , \begin{bmatrix} Z_j & I \\ I & Y_j \end{bmatrix} > 0 \tag{12}$$

for all $j \in I_m$.

Using distinct matrix bounds on the performance functions has the advantage of reducing conservatism in the solution but has the disadvantage of the loss of matrix constraint convexity. Such convexity is required for the application of the interior point algorithms. However the reduction in conservatism can be recovered to some extent with a decomposition of the nonconvex Optimization Problem 3.1 into a coupled set of convex problems. We consider the following Algorithm 3.2 to be a numerical analog of the alternating projections method of Grigoriadis and Skelton [6].

Algorithm 3.2

Step 1 Guess initial values of *the* $3m + 1$ *matrix optimization variables* K and $\{(X_j, Y_j, Z_j) : Y_j = Y_j^T > 0, Z_j = Z_j^T > 0\}_{j \in I_m}$ and a scalar variable t. Calculate fold= $tr\{Z_1\} + \ldots + tr\{Z_m\} + t$ *for* use *in* a convergence tolerance *test* in *Step 4*.

Step 2 Minimize *the* magnitude *(that is, the maximum* singular value) *of the* simultaneously *stabilizing gain* matrix K *for fixed values of* matrices $\{(X_j, Y_j, Z_j)\}_{j \in I_m}$. *Vandenberghe and Boyd [12] show that this minimization* can *be posed in terms of linear* matrix inequalities as

$$\min_{K,t} \; t$$

subject to the linear *matrix equality constraints*

$$X_j = K Y_j, \forall j \in I_m \tag{13}$$

for given matrix pairs $\{(X_j, Y_j)\}_{j\in I_m}$; and *subject to the* linear matrix inequality constraint

$$\begin{bmatrix} tI & K \\ K^T & tI \end{bmatrix} > 0 .$$

The converged values of variables K and t are then saved for use as fixed quantities in Step 3.

Step 3 Minimize *the guaranteed-cost bounds* for *each of the m systems by*

$$\min_{\{(Y_j,Z_j)\}_{j\in I_m}} \sum_{j=1}^{m} tr\{Z_j\}$$

subject to the $2m$ *linear* matrix *inequality constraints*

$$\begin{bmatrix} -Y_jA_j^T - A_j Y_j + B_jKY_j + Y_jK^TB_j^T & Y_j & Y_jK^T \\ Y_j & Q_j^{-1} & 0 \\ KY_j & 0 & R_j^{-1} \end{bmatrix} > 0 , \quad \begin{bmatrix} Z_j & I \\ I & Y_j \end{bmatrix} > 0$$

for given K from Step 2 for *all* $j \in I_m$. *Save the* converged *values of optimization variables* $\{Y_j\}_{j\in I_m}$ *for use as fixed quantities in a return to Step 2. Calculate* matrices $X_j = KY_j$ *for all* $j \in I_m$, also *for* use as *fixed* quantities in *Step* 2.

Step 4 *Check for convergence: let* $f_{new} = tr\{Z_1\} + \ldots + tr\{Z_m\} + t$. *If* $\|f_{new} - f_{old}\| < \varepsilon$ (some *user-defined tolerance), then stop. Otherwise let* $f_{old} := f_{new}$ and go *to Step 2.*

An immediate and obvious wrinkle is that constraints (13) specify an overdetermined system. For example, in the flight trajectory problem described in Section 2, constraints (13) involve four matrix equations (twelve scalar equations) in three scalar unknowns (three components of the gain K array). However it is well known[1] that linear matrix inequalities (13) admit at least one solution K if $\mathrm{rank}(Y^T) = \mathrm{rank}([Y^T \mid X^T])$ where $X := [X1, \ldots, X_m]$ and $Y := [Y_1, \ldots, Y_n]$. Analysis of the example problem using MATLAB indicates that the ranks are in fact equal. Further, we find that solution K is *unique* because Y^T has full rank.

It is already known that there is at least one solution, namely the "conservative" $K^* = X^*(Y^*)^{-1}$ from optimization Problem 2.1, assuming $X := X_1 = \cdots = X_m$; $Y := Y_1 = \ldots = Y_m$; and $Z := Z_1 = \ldots = Z_m$. Step 2 can thus be replaced entirely, obviating the need for iteration. Also since K is constant in Step 3, the sufficient condition matrix Lyapunov inequality (4) becomes, after replacing matrix variable P with distinct matrix variables $\{P_j\}_{j\in I_m}$,

$$(A_j - B_jK)^T P_j + P_j(A_j - B_jK) + Q_j + K^TR_jK < 0.$$

This is already a linear matrix inequality and Algorithm 3.2 is therefore simplified to the following Algorithm 3.3.

Algorithm 3.3

Step **1** *Solve the "conservative" (optimization Problem 2.1* and *calculate the resulting gain matrix K* as

$$K = XY^{-1} .$$

Save the converged value of K and use it as a fixed quantity in Step 2. Use the resulting conservatively *optimized values of X, Y, and Z as* initial guesses *of the matrix* variables $\{(X_j, Y_j, Z_j) : Y_j = Y_j^T > 0, Z_j = Z_j^T > 0\}_{j\in I_m}$, *respectively, in Step 2.*

---

[1] A solution x to an overdetermined system $Ax = b$ exists if $\mathrm{rank}(A) = \mathrm{rank}([A \mid Y])$. This is due to the fact that the equality of ranks implies that b is in the column space of A. The solution is unique if A *is* of full rank because that implies that the null space of A is of dimension zero. If A were not of full rank then there would exist a nonzero vector $z \neq x$ contained by the null space of A. Vector z would be a solution of $Az = b$ because $b = A(x + z) = Ax + Az$. Since z is in the null space of A, $Az = O$, leaving $Ax = b$.

Step 2 Minimize *the guaranteed-cost* bounds for each *of the m systems by*

$$\min_{\{(P_j\}_{j\in I_m}} \sum_{j=1}^{m} tr\{P_j\}$$

subject to the m linear matrix *inequality constraints*

$$-A_j^T P_j - P_j A_j + K^T B_j^T P_j + P_j B_j K - Q_j - K^T R_j K > 0$$

for all $j \in I_m$.

The example used above for the "conservative" Optimization Problem 2.1 is revisited for use with the noniterative Algorithm 3.3 of reduced conservatism so that results can be compared. The optimized matrix variables turn out to be

$$P_1^* = \begin{bmatrix} 0.0726 & -0.3125 & 0.1774 \\ -0.3125 & 1.8192 & -0.9305 \\ 0.1774 & -0.9305 & 0.5175 \end{bmatrix}, \quad P_2^* = \begin{bmatrix} 0.0479 & -0.5360 & 0.3625 \\ -0.5360 & 7.2761 & -4.6376 \\ 0.3625 & -4.6376 & 3.0744 \end{bmatrix},$$

$$P_3^* = \begin{bmatrix} 0.2593 & 1.6214 & 0.66481 \\ 1.6214 & 12.9840 & 3.3934 \\ 0.6648 & 3.3934 & 2.0399 \end{bmatrix}, \quad P_4^* = \begin{bmatrix} 0.1472 & 0.1014 & 0.80081 \\ 0.1014 & 0.5332 & 0.5158 \\ 0.8008 & 0.5158 & 4.5380 \end{bmatrix}.$$

The maximal guaranteed-cost bound over all m = 4 systems is calculated as

$$\max_{j\in I_4} \mathrm{tr}\left\{P_j^*\right\} = 15.3\,,$$

a reduction of an order of magnitude (at least for this example) over the conservative bound (9). Note that the ability to simultaneously control all m systems with the single gain matrix K is retained.

4 Summary

This paper demonstrates a decomposition of the convex linear-quadratic simultaneous performance design met hod that can produce a significant reduction in solution conservatism. Previous work by the authors is reviewed, numerical analysis and simplification of the decomposition is provided, and examples are given for the purpose of comparison.

References

[1] J. Bernussou, P.L.D. Peres, and J.C. Geromel. A Linear Programming Oriented Procedure for Quadratic Stabilization of Uncertain Systems. *Systems & Control Letters,* 13(1):65–72, July 1989.

[2] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. Linear Matrix *Inequalities in System and* Control Theory. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1994.

[3] S.S.L. Chang and T.K.C. Peng. Adaptive Guaranteed Cost Control of Systems with Uncertain Parameters. *IEEE* Transactions on *Automatic Control,* AC-17(4):474–483, August 1972.

[4] P. Dorato, C.T. Abdallah, and V. Cerone. *Linear* Quadratic *Control:* An Introduction. Prentice-Hall, Englewood Cliffs, New Jersey, 1995.

[5] L. El Ghaoui, F. Delebeque, and R. Nikoukhah. LMITOOL: A User-Friendly Interface for LMI Optimization. Preprint downloaded from WWW server at Ecole Nationale Supérieure de Techniques Avançées (ENSTA), **France,** February 1995.

[6] K.M. Grigoriadis and R.E. Skelton. Low-Order Control Design for LMI Problems Using Alternating Projection Methods. Automatic, 32(8), August 1996.

[7] I.N. Kar. *Parametric Conditions for* Simultaneous *Quadratic Stabilization of* Linear Systems. PhD thesis, Indian Institute of Technology Kanpur, India, May 1996.

[8] R.A. Luke, P. Dorato, and C.T. Abdallah. Linear-Quadratic Simultaneous Performance Design. Submitted to American Control Conference, Albuquerque, NM. Preprint available at http://www.eece.unm.edu/students/luke/, September 1996.

[9] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in* Convex Programming, volume 13 of *SIAM Studies in Applied Mathematics.* Society for Industrial and Applied Mathematics, Philadelphia, 1994.

[10] M. Paskota, V. Sreeram, K.L. Tee, and A.I. Mees. Optimal Simultaneous Stabilization of Linear Single-Input Systems Via Linear State Feedback Control. *International Journal of* Control, 60(4):483–498, 1994.

[11] L. Vandenberghe and S. Boyd. *SP: Software for Semidefinite* Programming. Stanford University Information Sciences Laboratory, Stanford, California, beta version edition, November 1994. User's Guide downloaded from anonymous ftp server at isl.stanford.edu in directory /pub/boyd/semidef_prog/.

[12] L. Vandenberghe and S. Boyd. Semidefinite Programming. *SIAM Review,* 38(1), March 1996.

URC97083

An Extremely Slow Motion Servo Control Technique used in the Hubble Space Telescope's Star Selector Servo Subsystem

Asad M. Madni[1] and Mike Jumper[2]

[1]BEI Sensors& Systems Company
13100 Telfair Avenue
Sylmar, CA 91342
Tel: (818) 364-7215
Fax: (818) 362-1836
E-Mail: bei1madni@aol.com

[2]BEI Sensors& Systems Company
Precision Systems & Space Division
1100 Murphy Drive
Maumelle, AR 72113
Tel: (501) 851-4000
Fax: (501) 851-5476

Abstract

A state-of-the-art, slow motion dual axis servo control system used on the Hubble Space Telescope's Star Selector Subsystem is presented. It utilizes optical encoders with integral DC torque motors to provide precise, digital rate control over a range of 0.5 arcsecond/second (0.0333 revolution/day) to 16,384 arcseconds/second (0.759 revolution/minute) and 21-bit absolute-position words to an accuracy of 2.0 arcseconds peak-to-peak. In the fine-tracking mode the system provides position tracking to the order of 0.1 microradian.

Introduction

In order to accomplish extremely precise pointing and stabilizing, the Hubble Space Telescope [1,2] uses a combination of rate gyros, reaction wheels, star trackers and interferometric fine guidance sensors. The gyros provide a reference frame so that the Space Telescope "knows" its initial orientation and the position to which it must be moved. On command from the ground, the reaction wheels (23 inches or 59 cm in diameter) are accelerated or decelerated to slew the Space Telescope to a new object. Star trackers then use known pre-selected bright stars to ascertain the pointing direction to within about half an arc-minute (one-sixtieth of the apparent diameter of the full moon). Two of the three interferometric fine guidance sensors begin to function when a guide star has been brought within an area covering about 60 square arc-minutes (about one-tenth of the moon's apparent area). The fine guidance sensors, acting through the rate gyros and the reaction wheels, point the Space Telescope at the target to within 0.02 arc-second. This is like pointing at the face of a U.S. quarter dollar as seen from 150 miles (or 240 Kilometers) away. In addition to pointing accuracy, pointing stability is an equally vital critical factor. Over a 24 hour period, the pointing direction must not waver by as much as that quarter's width. This stability is necessary for the Space Telescope to collect and integrate light from very distant target objects for up to 10 hours. (For 14 out of the 24 hours, the Sun, Moon, Earth or Van Allen radiation belts can interfere with data-taking and interrupt the continuous integration of target light).

The Star Selector Servo Subsystem is a part of the Fine Guidance System (FGS) of the Space Telescope. It is used in conjunction with the fine guidance computer to locate and position lock onto particular stars. In order to find the star of interest, the Star Selector Servo Subsystem is either scanned or commanded to particular position by the FGS through use of the velocity inputs using the position feedback from encoders to ascertain pointing direction. Once the star is within the capture range of the fine guidance sensor, the output of the FGS is used for feedback control. The near continuous position control of the Star Selector Servo Subsystem allows a fine lock to the FGS to within 0.1 microradians, thus providing the highly stable reference required for pointing the telescope.

Conventional servo systems utilizes a DC or digital tachometer as the feedback device for comparison with a reference signal with the resultant error signal being used to drive the motor. A major limitation of this approach is that the feedback signal gets so slow or so erratic as to become unusable at extremely slow or stopped conditions. To overcome this limitation the Star Selector Servo Subsystem uses the ULTRA-LOC® technique [3,4] which is a carrier-based servo system utilizing a signal that is the result of an encoder interpolation technique known as the

Optical Resolver®. With this interpolation technique, a logic-level square signal, referred to as a Phase Variable (PV) is generated such that its phase relationship to the input carrier signal ($\omega_c t$) is a linear function of the shaft position (θ). The PV signal forms the basis for both high-resolution encoding and the ability to provide ultra-smooth rotational control at incredibly slow sidereal speeds. The following sections provide a mathematical analysis of the ULTRA-LOC® technique and a functional description of the Star Selector Servo Subsystem.

MATHEMATICAL REPRESENTATION OF SYSTEM OPERATION

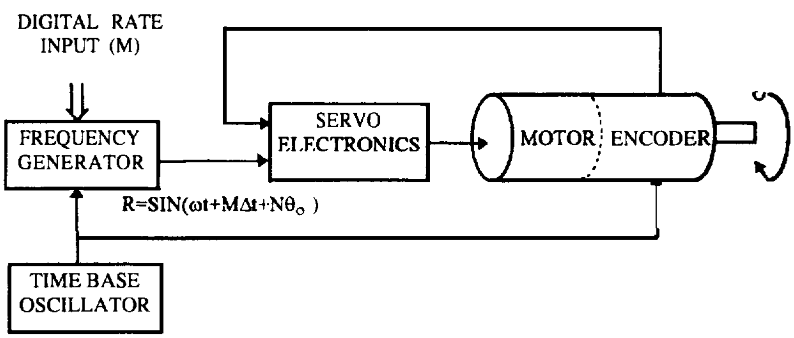


Figure 1. ULTRA-LOC® Rate System Block Diagram

As mentioned previously, the ULTRA-LOC® technique uses a carrier-based optical resolver to create a continuous feedback signal even if rotation is stopped. The signal is similar to the signal received from magnetic resolvers except that carrier frequencies and resolutions are several times higher [5].

The master oscillator in Figure 1 drives the carrier frequency generator which creates the signals Sin $\omega_c t$ and Cos $\omega_c t$ which are used to modulate the optical resolver. The resovler disk contains N cycles of each sine and cosine waveforms which, when the shaft is rotated, create the signals:

$$\text{Sin } (N\ d\Theta/dt) = \text{Sin } (N\ \omega_\theta t) \text{ and Cos } (N\ d\Theta/dt) = \text{Cos } (N\ \omega_\theta t) \qquad (1)$$

where ω_θ is the angular speed of shaft rotation. If Sin ($N\ \omega_\theta t$) is used to modulate Cos ($\omega_c t$) and Cos ($N\ \omega_\theta t$) is used to modulate Sin ($\omega_c t$) and the two modulated signals are added together, the resulting signal may be described by the following expression:

$$\text{Sin } (N\ \omega_\theta t) \text{ Cos } (\omega_c t) + \text{Cos } (N\ \omega_\theta t) \text{ Sin } (\omega_c t) = \text{Sin } (\omega_c + N\ \omega_\theta)t = \text{Sin } 211\ (f_c + Nf_\theta) \qquad (2)$$

where f_c = carrier frequency and f_θ = rotational frequency. The master oscillator is also used to drive a digitally controlled frequency synthesizer whose output is defined by:

$$f_0 = f_c\ [\ 1 + (M/2^{18})] \qquad (3)$$

where M is an integer obtained from the computer data bus with a range of: $-(2^{15}) \le M \le (2^{15} - 1)$ and includes O. The outputs of the resolver and the synthesizer arc fed to a phase comparator that generates an analog error signal for frequency and phase differences between the two signals. This error signal after amplification and compensation is used to drive the motor until a null is reached. If the two signals are of equal frequency then:

$$f_0 = f_c\ [\ 1 + (M/2^{18})] = f_c + Nf_\theta \text{ or } f_c + (Mf_c\ /2^{18}) = f_c + Nf_\theta \qquad (4)$$

$$Nf_\theta = Mf_c\ /2^{18} \text{ and } f_\theta = [\ M/N\ (2^{18})]\ f_c$$

where: f_θ = rotational frequency (Rev/see), f_c = carrier frequency (Hz), N = number cycles sine/cosine contained on disk (cycle/Rev) and M = speed multiplier variable from computer.

For an example of a speed control system, let f_c = 32.768KHz (2^{15} Hz) and N = 2048 cycles/rev. ($2^{11} \cong$ /rev), then f_θ = [(2^{15})M / (2”) (218)] = M / 2^{14} Rev/see, since - (2^{15}) ≤ M ≤ (2^{15} - 1), then the speed range is for M = (2^{15}), f_θ = $-2^{15}/\ 2^{14}$ = -2 Rev/see (maximum reverse speed), for M = (2^{15} - 1) f_θ = 2*^{5}-1 / 2*^{4} = 32,767 /16,384 = 1.999387 Rev/see (maximum forward speed), then slowest non-zero speed is (M = 1) f_θ = 1 / (2^{14}) = 2^{-14} 0,000061035 Rev/see or 0.0036621 RPM.

This represents the step size by which the speed maybe incremented. Various combinations of f_c and N maybe used to determine the required speed range or step size. N maybe any number, but powers of two are commonly used in order to extend the resolution of the absolute binary position word as well as provide speed control [6].

System Description

The Star Selector Servo Subsystem is a modular, dual-axis servo system consisting of two encoder-motor assemblies, two remote electronic units and one dual servo unit (see Figure 2). The dual servo unit is 15.06 cm x 22.86 cm x 15.39 cm H, the remote unit is 21.59 cm x 17.78 cm x 12.85 cm H the encoder-motor is 27.94 cm diameter x 15.06 cm H (excluding mounting feed and cable protrusions). All units are designed for conduction cooling as required for space-borne equipment. The mechanical parts of the encoder-motor assemblies are fabricated from beryllium; each assembly has a mass of approximately 9.55 Kg. The electronic boxes are made of 6061 -T6 aluminum for mass of 3.18 Kg for the remote unit and 5.9 Kg for the servo unit.

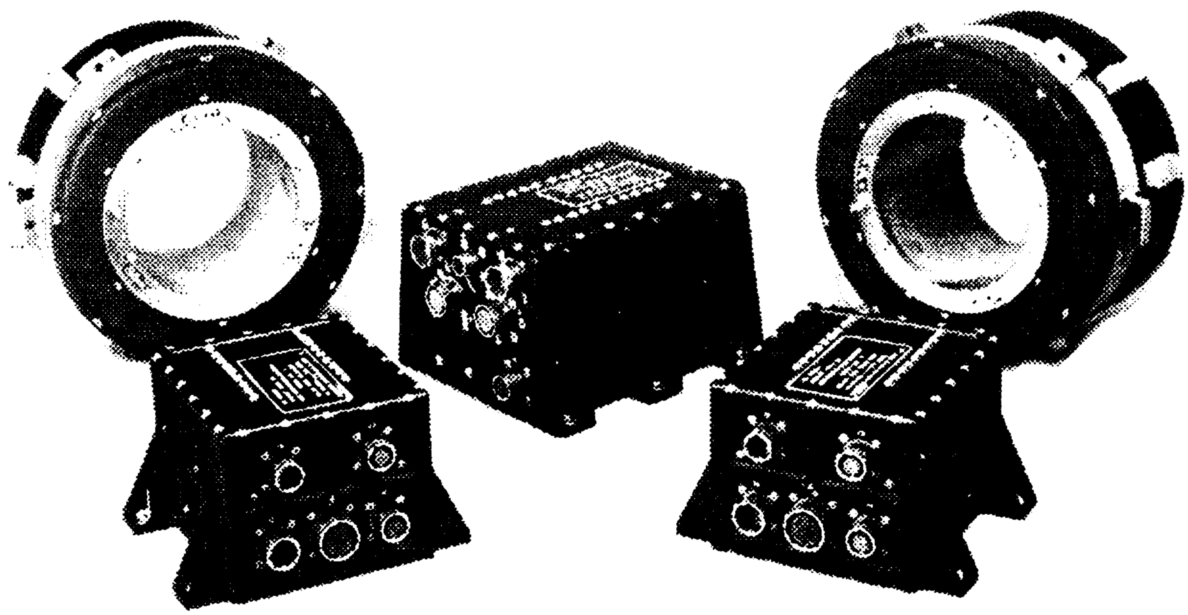


Figure 2. Star Selector Servo Subsystem

The system accepts two 16-bit velocity words along with motor on/off and controls. The output is two 21 -bit binary position words. As shown in the block diagram (Figure 3), the servo unit provides the reference signal, the basic motor control function and a one's complementor circuit for reversing the encoder output word. The dual servo unit also provides electrical interface for the subsystem. The remote units provide encoder translation electronics and the servo loop control. Each encoder-motor assembly provides position readout and motor function, motor commutation, signal preamplification and motor servo power amplifier functions.

Encoder-Motor Assembly

The encoder-motor assembly is a key element of this subsystem. The encoder provides 15 bits of Gray code for "coarse" information a set of sine/cosine outputs encoded at 16,384 cycles/revolution for "fine" information and a set of sine/cosine tracks encoded at 12 cycles/revolution for motor commutation. The gray code data is obtained from a single read station utilizing an edge-emitting LED pulsed at 3.0 Amperes for approximately 15 microseconds. The light passing through the Gray code tracks and their complements is detected by a 30-cell monolithic photocell array. Preamplifier and comparator circuits contained in the encoder transform the photocell signals into logic level pulses.

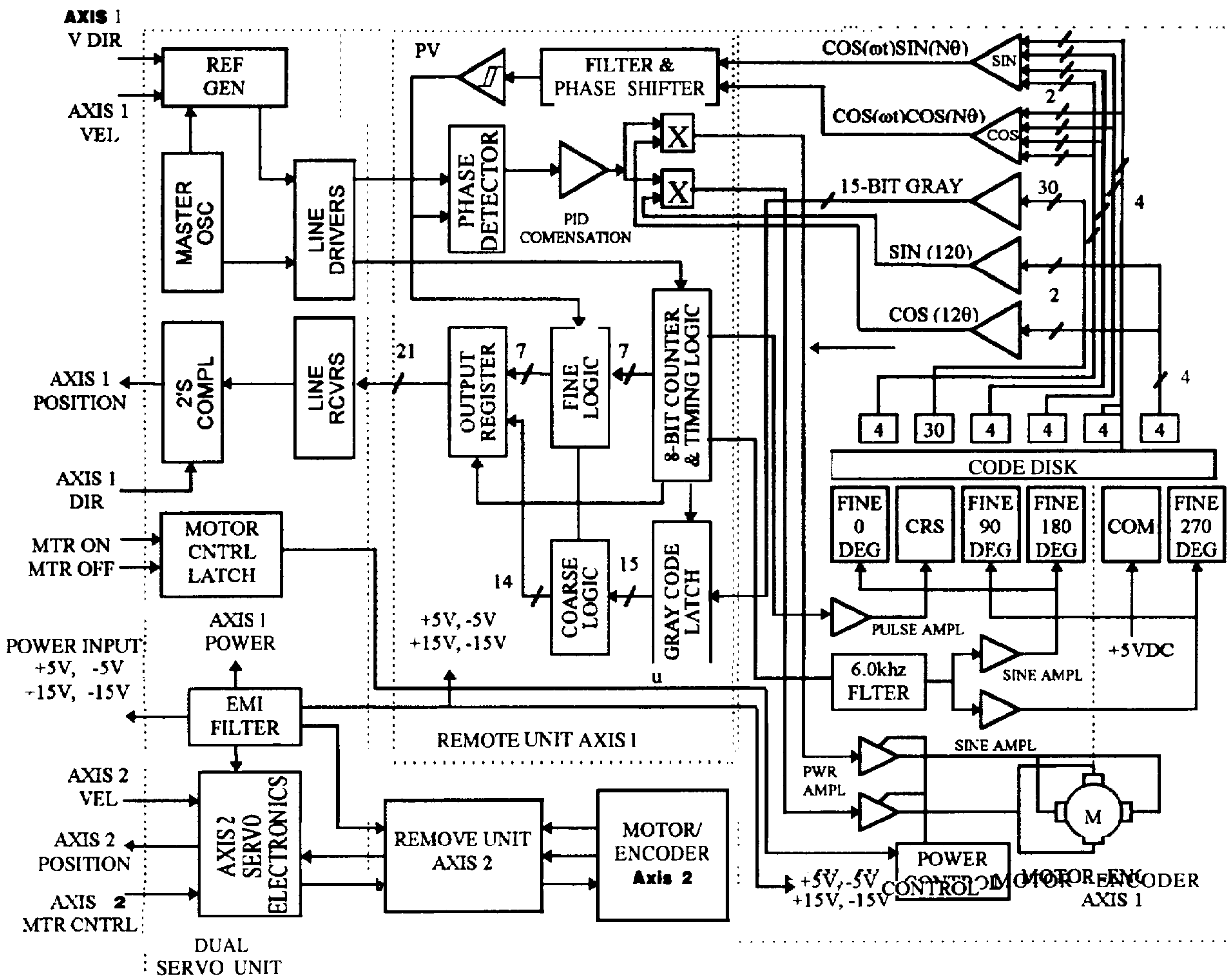


Figure 3. Star Selector Servo Subsystem Block Diagram

The fine tracks are used to extend the encoder word to 21-bit resolution using the Optical Resolver interpolation technique. This method of interpolation allows the sine and cosine tracks to modulate an input carrier to obtain signals similar to those obtained from a magnetic resolver. To accomplish this, the four fine LEDs are driven with a 6,628 KHz sine wave carrier signal. Two voltage-to-current amplifiers driven from a three-pale active filter provide this function. Four optic stations are used to cancel errors that might be introduced by bearing runout, variation in disk flatness or variations in code pattern. One count on a 21-bit encoder represents 0.618 arcsecond or 2.94 microradians on a 25.4 cm dia. code disk,

The signals from the four fine read stations arc preamplifier and summed within the encoder to provide the following calibrated signals:

$$A = \cos(\omega_\theta t)\,\sin(N\theta) \text{ and } B = \sin(\omega_c t)\,\cos(N\theta) \qquad (5)$$

These two signals are further processed in the remote unit to accomplish the extended resolution and provide the servo feedback function. The two sinusoidal motor commutation signals are amplified through calibrated preamplifiers in order to provide a low ripple torque signal to a two-phase, 24-pole brushless DC torque motor. The encoder- motor assemblies also contain two drive amplifiers for driving the motor.

Dual Servo and Remote Units

Each remote unit provides the subsystem with encoder translation and servo loop control, while the dual servo unit

provides electrical interface and the input control functions for the servo. The functionality of the dual servo and remote units are intertwined and therefore discussed concurrently in the following sections.

Analog Resolver Section

The two modulated carrier signals from the encoder are passed through a phase shifter and filter section to obtain the signals Sin $(\omega_c t)$ Sin $(N\theta)$. These two signals are combined according to the trigonometric identity Sin X Cos Y - Cos X Sin Y = Sin (X-Y) to yield the function:

$$C = \text{Sin}\,(\omega_c t - N\theta) \tag{6}$$

This signal is transformed into a logic level square wave signal by means of a zero crossing comparator. This square wave signal is referred to as the Phase Variable (PV) because the phase position of its leading edge, relative to the leading edge of the input carrier square wave, represents the fine position of the shaft. The carrier signal is derived from the MSB of a divide by 256 counter driven from the master oscillator. The phase of the PV signal relative to the counter MSB is proportional to the shaft angle. This phase angle is measured by latching the 8-bit count at the time the PV signal transitions from low to high. A digital offset is then added to the digitized position to align it to the coarse or Gray code data.

Coarse data is stored in a 15-bit register in Gray code form. The data word is then transformed into natural binary and passed to a binary adder for merging with the fine data. After the data is translated and merged, the resultant 2 l-bit encoder word is stored in the output storage register. This occurs approximately 10 microseconds after the leading edge of PV allowing time for the data to propagate through the merging logic. A marker pulse is provided at each data update so that data transfer during output transitions maybe avoided. The 2 l-bit encoder data word is buffered by line drivers and passed over a 10.7 meter paired cable to the dual servo unit. A one's complement circuit in the servo unit provides direction control of the encoder count.

Servo Loop Control

As previously mentioned, the PV signal serves as the feedback for servo loop control (see Figure 4). This signal is phase locked to a synthesized square wave generated in the dual servo unit which has the output frequency:

$$f_R = f_c\,(1 + M / 2^{23}) \tag{7}$$

where f_R = reference frequency, f_c = carrier frequency and M = 16-bit signal velocity number. Both the reference signal and the carrier are obtained from the master oscillator contained in the dual servo unit to make the basic carrier reference exactly the same in both units.

Since the phase of the PV signal in relation to the fixed carrier is equal to $-N\theta$, the frequency of PV will be shifted as the shaft is turned. It will shift up if the shaft is turned in the counter-clockwise direction and down for clockwise rotation. If the encoder PV output frequency is considered with the shaft turning, it would be shown as:

$$f_{pv} = f_c + Nf_s \tag{8}$$

where : f_s = shaft frequency and N = number of Sine/Cosine cycle on the disk (16,384). If this signal is put into frequency and phase lock with the signal from the reference generator such that $f_{pv} = f_R$, then:

$$f_c + Nf_s = f_c\,(1 + M/2^{20}) \qquad \text{where: } Nf_s = f_c\,M/2^{20} \quad \text{and} \quad f_s = (M/2^{20})\,f_c \tag{9}$$

The rotation of the shaft is now locked to the velocity word input of the reference generator, even if the velocity word is zero. Furthermore, the shaft velocity is directly proportional to the carrier frequency. Since a crystal-controlled oscillator is used to generate the carrier frequency and to drive the synthesizer, the average rotational speed of the shaft is to the same precision as that of the oscillator.

It should be noted that the frequency synthesizer uses a 13-bit binary word to generate a triangular analog waveform. This waveform is smoothed to a near Sine wave and then zero-detected to produce the square wave

output. This makes 2^{13} distinct phases relative to the fixed carrier available from the reference generator. The encoder has 16,384 cycles encoded on its disk so that one complete cycle of phase change represents one Sine/Cosine cycle or 79.1 arcseconds. This gives the reference generator a position resolution of 79. $1/2^{13}$ or 0.00966 arcsecond (0.0468 micro-radian). Note that this is 64 times more resolution than the 21-bit encoder word or 0.0058 micrometers of rotation on the 25.4cm track diameter [7].

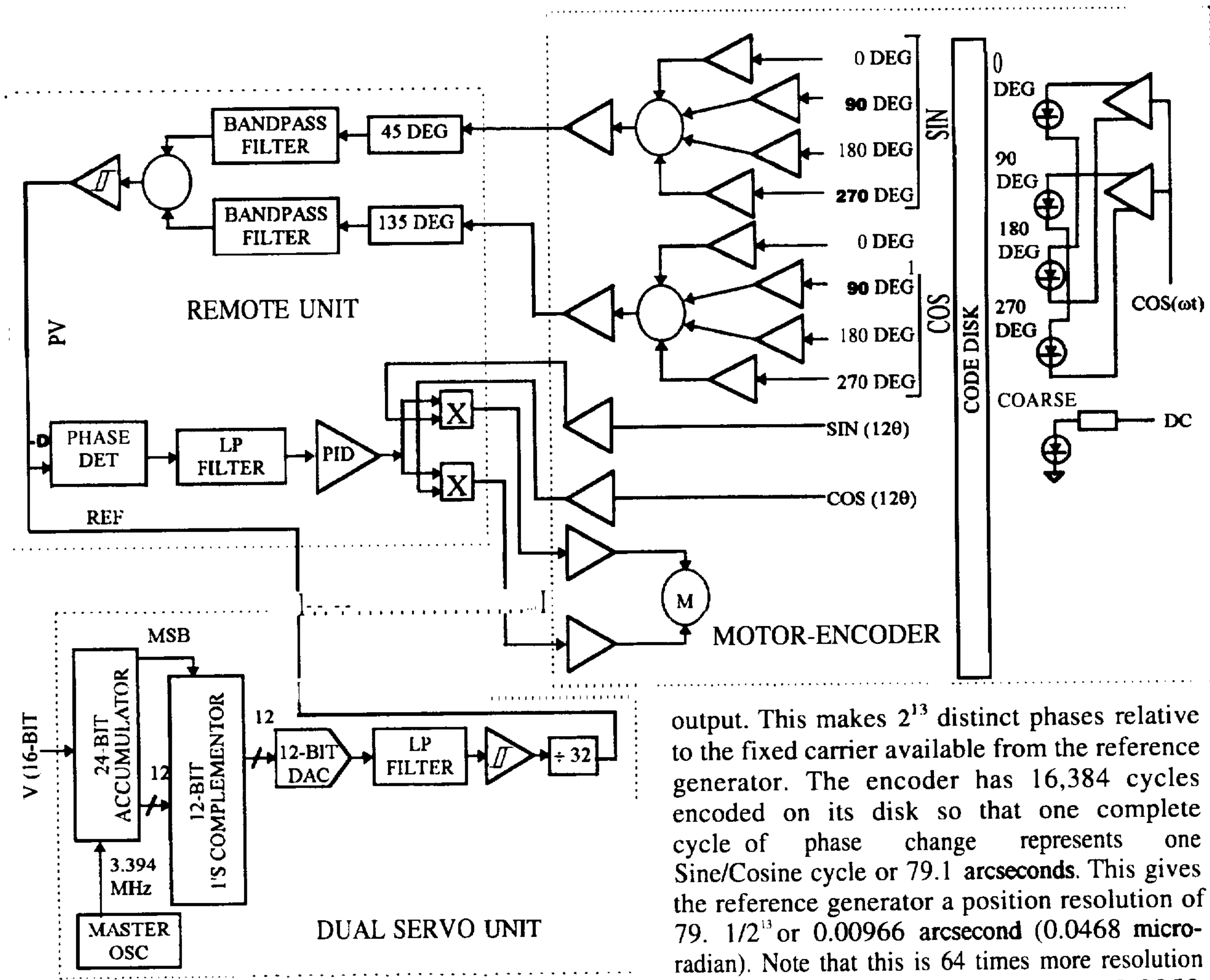


Figure 4. Servo Loop Block Diagram

The carrier frequency selected for the Star Selector Servo Subsystem is 6,62804 KHz. This selection was made to obtain the desired units of velocity (0.5 arcsecond/sec. steps) and to maintain a low delta f_{PV} (± 207 Hz) at the maximum velocity. This carrier frequency maybe verified by application of equation mentioned above.

The servo loop incorporates a three-state phase detector followed by a low pass smoothing filter and PID compensated amplifier section. The compensated error signal is then passed through analog multipliers to combine it with the two commutation phases. The two-phase output is then passed to the two motor amplifiers contained in the encoder-motor assembly. This is a type 111 servo loop having a -3dB point at approximately 150 Hz and a positional gain of 1.625 oz-in/arc-sec.

The non-linear bearing friction characteristics observed over small angles [8,9] were approximated in the frequency domain model by a linear spring constant (K_S) and a viscous friction damping constant (K_V). The rotor load transfer function used was:

$$\frac{1}{JS^2 + K_V\, S + K_S} \tag{10}$$

Both the viscous friction constant (K_V) and the spring constant (K_S) were determined by test. Verification of spring constant occurred during open loop frequency tests which show a natural frequency at:

$$\frac{1}{2\pi}\sqrt{\frac{K_S}{J}} \tag{11}$$

The viscous friction constant was determined by measuring the torque differential between the maximum and minimum velocity.

System Performance

Five flight-ready systems (including the qualification model) were delivered, with performance equal to or exceeding specifications. The Hubble Space Telescope was launched in April 1990, with its Star Selector Servo Subsystem on-orbit performance exceeding specifications to date. The accuracy of the encoders was characterized for both wide and narrow angle errors and is listed in Table I. The tests were conducted using a Moore Model 1440 Divided Engine and a Hewlett-Packard Model 5500 Laser Interferometer. The results are accurate to 0.2 arcsecond and 0.1 arcsecond for the wide angle and narrow angle measurements respectively.

Table I. Encoder Accuracy

Encoder-Motor Serial No.	Narrow Angle Peak-to-Peak (arc-see)	Wide Angle Peak-to-Peak (arc-see)	Total Peak-to-Peak (arc-see)	Total RSS (arc-see)
M/E 3005	0.57	1.1	1.67	0.35
M/E 3006	1.11	0.9	2.01	0.45
M/E 3007	0.80	1.15	1.95	0.43
M/E 3008	0.78	1.0	1.78	0.36
M/E 3009	0.71	0.80	1.51	0.28
M/E 3010	0.80	14	2.2	0.52

Independent verification of the rate response to a minimum command input was done by Lockheed Missiles and Space Co., Inc. in Sunnyvale, California, [10] and the final report stated that: "The conclusions reached from these tests were: (1) The ULTRA-LOC® Star Selector Servos do respond to LSB level (0.5 arcsecond/sec. rate commands), (2) The response to small rate commands can be measured even when the net position change is less than one encoder LSB (0.618 arcseconds), (3) Star Selector Servo response to small rate commands is (on average) proportional to commands, despite the fact that the level is within the Dahl Friction characteristic range".

CONCLUSION

The design for a state-of-the-art, extremely slow motion dual axis servo control system used on the Hubble Space Telescope's Star Selector Subsystem is presented. The control system is a closed-loop system that provides position or velocity control of a rotating shaft or table with the stability and accuracy required by the most demanding applications. It accepts rate or position inputs from computer I/O's and provides shaft outputs that are precisely proportional to either a rotational rate or a position. The technique can be adapted to numerous applications including: rate tables, one and two axis tracking, indexing (repeated incremental positioning), Z-axis pointing and positioning (of telescopes, lasers, weapons, mirrors and other sensors), motion translation (rotary to linear or linear to rotary), inertia compensations, micromanipulation, electronic gearing, force feedback, coordinate transformation and multiple slaves (precision line shaft drives).

ACKNOWLEDGMENT

The authors wish to acknowledge Dr. Mohammad "Mo" Jamshidi, Director, NASA ACE Center and AT&T Professor, University of New Mexico, Albuquerque for providing the enthusiasm in publishing this work on the Hubble Space Telescope. The authors also wish to express their gratitude to Linet Aghassi for her patient and meticulous typing of the manuscript.

REFERENCES

[1] "Hubble Space Telescope", Publication by National Aeronautics and Space Administration (NASA), Marshall Space Flight Center.

[2] "Hubble Space Telescope Fact Sheet", Publication by Perkin-Elmer Space Telescope Program Office, 100 Wooster Heights Road, Danbury, CT, 06810.

[3] Hafle, Ralph S., 'Wide Range, High Accuracy Electronically Programmed Speed Control System", U.S. Patent No. 3,974,428, issued August 10, 1976. Assignee: Baldwin Electronics, Inc., Little Rock, AR.

[4] Hafle, Ralph S., "Geschwindigkeits - Steverungs - System", Deutschland Patent No. DE2521355C3, issued July 6, 1978. Assignee: BEI Electronics, Inc. Little Rock, AR.

[5] Boycs, G. S., "Synchro and Resolver Conversion", Memory Devices, Ltd., Surrey, United Kingdom, 1980.

[6] Hafle, Ralph S., and Cole, C. F., "A Wide Range, High Accuracy, Electronially Programmable Ultra-Loc® Speed Control System", BEI Electronics, Inc., Maumelle, AR, 1979.

[7] "181 Series Servo Controller", Operators Manual, BEI Electronics, Inc., Maumelle, AR.

[8] Dahl, P. R., "A Solid Friction Model", The Aerospace Corporation, El Segundo, CA., 1%8

[9] Perry, J., "Small Angle Behavior of Bearing Friction Torque", BEI Electronics, Inc., Maumelle, AR.

[10] "Star Selector Servo Response to Minimum Command Input Bit", Lockheed Missiles and Space Co., Inc., Sunnyvale, CA., ST/SE-24, Section G, Part 19, 1986.

BIOGRAPHIES

Dr. Asad M. Madni is President and CEO of BEI Sensors and Systems Company, Inc. located in Sylmar, California. Prior to joining BEI in 1992, he was with Systron Dormer Corporation (A Thorn/EMI Company) for 18 years where he served in various senior level technical and managerial positions, eventually as Chairman, President and CEO. He received the A.A.S. degree from RCA Institutes, Inc., B.S. and M.S. degrees from [University of California, Los Angeles (UCLA) and the Ph.D. degree from California Coast University (CCU), all in electrical engineering. He is also a graduate of the Certificate Program in Engineering Management at the California Institute of Technology, the Executive Institute at Stanford University and the Program for Senior Executives at MIT Sloan School of Management. Dr. Madni is an internationally recognized authority with over 25 years of experience in the field of intelligent system design and signal processing. He is credited with over 60 publications and numerous U.S. patents resulting in "industry firsts". He serves in an advisory capacity for several professional and academic organizations. Dr. Madni is also the recipient of numerous awards and honors and is listed in 15 Who's Who publications including Who's Who in America. He is a Senior Member of the Institute of Electrical and Electronics Engineers (IEEE), Life Member of the Association of Old Crows (AOC) and Member of the American Association for the Advancement of Science (%4,4 S), New York Academy of Sciences (NYAS) and the Society of Automotive Engineers (SAE).

Mike Jumper is presently serving in the capacity of Business Development Manager for BEI Sensors and Systems Company's Precision Systems and Space Division. He was the System Mechanical Design Engineering for the Hubble Star Selector Servo Subsystem during the engineering and production programs. His responsibilities at BEI have included work on the GOES Scanning Mirror Motor/Encoders, UARS High Gain Antenna Position Sensors and the development of the 24 and 25-bit high accuracy commercial optical encoders. He received a BS degree in Mechanical Engineering from the University of Arkansas in 1980 prior to joining BEI that same year.

URC97084

Object-Oriented Design of a Drawing Subsystem For a Computational Mechanics Toolkit.

Anish Malanthara * Malcolm Panthaki t Walter Gerstle ‡ Raikanta Sahu §

November 18, 1996

Abstract

This paper describes the design and implementation of a device-independent and rendering-pipeline independent drawing subsystem for a Computational Mechanics Toolkit (Co MeT). CoMeT is an analysis environment for computational mechanics problems.

CoMeT is being designed and implemented in an object-oriented manner, C++ being the language used for implement ation. Objects (or classes, in a more general manner) that function closely together are organised as s *ubsyst ems* and subsystems that interact closely are grouped as primary *components*.

The Drawing subsystem is a group of classes that work closely together to render objects in CoMeT (CoMeT Data Objects or CDOs) on the screen, and is one of several subsystems that consolidate the *Job* primary component.

The Drawing subsystem has been designed based on the concept of a Scene. The Scene contains displayable objects. There are *GraphicsAttributes* associated with the each of the CDOs as well as with the entire Scene. The Scene interacts with the *GUI* subsystem in the *Command Source* primary component as well as third party graphics libraries through separate classes which handle this interaction to actually display the objects on the screen.

The presented design is one that can be used for any software which needs to display objects using a third party graphics library. The Drawing subsystem serves as a bridge between the rest of the code and the graphics library. Since only the Drawing subsystem knows about the graphics library, changing the underlying graphics library may be done with minimal effort and with minimal disruption to the rest of the code.

1 Introduction and Software Principles Behind the Design

1.1 CoMeT - a Computational Mechanics Toolkit

CoMeT is an object-oriented computational mechanics toolkit designed and implemented in Scheme and C++ programming languages. It is an analysis environment for computational mechanics problems. Geometric computations are handled by ACIS geometric modeling package [Spatial Technology Inc., 1996]. Other third party libraries such as *OpenInventor* [Silicon Graphics Inc., 1996] are also used in the overall implementation of CoMeT.

1.2 Overview of the Architecture of CoMeT

The atomic unit of the first level of abstraction in the architecture of CoMeT is called a *Primary Component* (PC) . This level is the highest level view of a complex system and facilitates ease of understanding as well as the first level of modularity for ease of implementation. A PC also serves as the unit of distribution of tasks across a network of computers (e.g. in a UNIX environment the PCs of CoMeT could conceivably be separate processes). Figure 1 shows the PCs of CoMeT. A connection between two PCs indicates that communication can occur between those PCs. By design, data redundancy across PCs is minimized.

*Graduate student, Dept. of Civil Engineering, University of New Mexico
t Staff, Dept. of Civil Engineering, University of New Mexico
‡ Associate professor, Dept. of Civil Engineering, University of New Mexico
§ Graduate student, Dept. of Civil Engineering, University of New Mexico

The *CommandSource* PC generates commands either through the GUI (Graphical User Interface) or a script file. This allows execution of CoMeT to be controlled either interactively (by user input through the GUI) or in batch mode (using a script file). Regardless of the source of the command, the *CommandSource* routes it to the Session PC. The Scheme functional programming language is used to represent commands in CoMeT.

A session (run) of CoMeT can consist of one or more users creating *Jobs* running concurrently. At any given time there is only one current *Job*. Commands from *CommandSource* are routed by the Session PC to the current Job. Each command is executed by a *Job*. As each *Job* is a separate process, the existence of multiple *Jobs* in a session allows for parallel execution of multiple sequential streams of commands, one per *Job*.

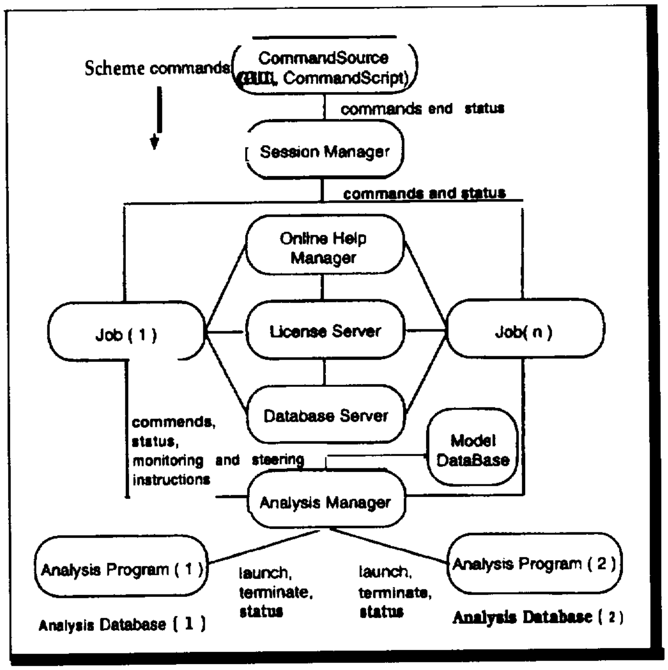


Figure 1: Architecture of CoMeT

The *DataBaseServer* PC provides access to and modification of model data. Multiple *Jobs* can access and/or modify this data simultaneously-the *DataBaseServer* will provide the necessary transaction management to prevent conflicts and the corruption of data.

The *Drawing Subsystem is a* subsystem in the *Job* primary component. It is responsible for displaying objects on the screen. In addition the *Job* primary component contains several subsystems that define and control how a problem is modeled, analyzed and visualized in CoMeT.

The Drawing subsystem is made up of several classes. These classes store and manipulate the data required by the Drawing subsystem. Several of these classes exist on a one-per-Job basis. Such classes are called *singletons* [Gamma et al., 1995].

1.3 Advantages of the Object-Oriented Approach to Software Design

In the real world, problems and solutions to them are composed of entities that have a certain amount of knowledge about the problem, and have certain functions they can perform with that knowledge. The entities involved in the solution often have knowledge that contributes the complete solution, and are expected to share that knowledge with other entities.

In software engineering, this means that the software design can correspond closely with the problem in the real world. Objects can correspond to the real entities and knowledge can be maintained within them. In practical terms, this means that the designers of the solution can have a greater understanding of the solution as it is designed. It also makes it easier to maintain and reuse the software for other similar problems with as few modifications as possible [Somerville, 1996] .

Inheritance and data encapsulation are two features of object-oriented design. Inheritance allows new objects to be derived from older ones, as extensions, often adding new features. Data encapsulation hides the data and implementation of procedures associated with an object from the user of that object. This allows the designer to change the algorithms or procedures used to perform a certain function in an object without affecting the way others use the object.

2 Design of the Drawing Subsystem

2.1 Scene

The Drawing subsystem has been designed around the concept of creating and manipulating Scenes. Much as in a real world view, a Scene contains *Drawable* objects, each of which has associated *GraphicsAttributes* and *GraphicsModes*. There are also *GraphicsAttributes* and *GraphicsModes* and *ViewParameters* and *LightingParameters* associated with the entire Scene, controlling the way the whole Scene is rendered. Each displayable object, or an item in the Drawable Object List (DOL) corresponds to a CometDataObject (CDO). Apart

from a pointer to the CDO itself, each item in the DOL also contains data required for rendering that particular CDO, such as *GraphicsAttributes* and *GraphicsModes*.

2.2 Use of Object-Oriented Design (OOD) and Data Encapsulation in the Drawing Subsystem

In Fig. 2, the classes in the Drawing Subsystem are shown, with hierarchies in a tree-like manner where they exist.

The actual rendering of *Drawable* objects is done using a third party graphics library. It is desirable at the design stage itself to isolate the graphics library as much as possible from other components. The design strives to isolate the knowledge of the graphics library to the objects which map the *Drawable* objects to the appropriate graphics library objects, and the objects which translate the CDO to be rendered into the set of graphics parameters and primitives that the graphics library can understand. In the present implementation of CoMeT, *OpenInventor* (OIV) [Silicon Graphics Inc. 1996] is the library utilised for rendering.

The main advantage of this choice of classes is the clean manner in which data and functionality gets divided between the objects. In the entire Drawing subsystem, *RenderingTools* and the *OIVNode Table* are the only classes which know about the third party library being used for graphics. This means that if a different library is to be used, the changes that need to be made are contained to these classes. Any information about the solid modeler used is cent ained in the *Rendering Tool* classes. With this design, the Drawing subsystem becomes as independent as possible from the underlying solid modeler and graphics library. This makes it easy to change the code to use the most effective third party library in the above mentioned categories.

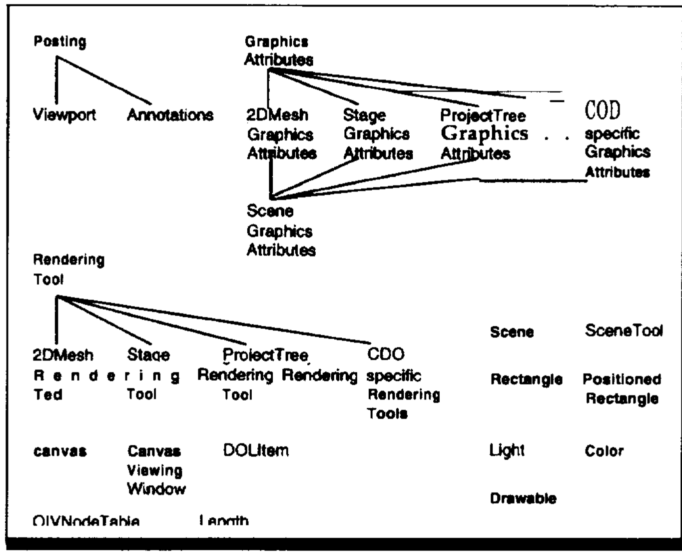


Figure 2: Chart of Objects in Drawing Subsystem, Showing Hierarchies where they exist

Adding another type of CDO to the existing type of displayable classes is also done easily. Graphicsization (generating graphics primitives for rendering) of a new CDO requires only the creation of new CDO specific *GraphicsAttributes* and *RenderingTool* classes. The CDOs themselves are kept free of any graphics functionality. The CDO-specific *Rendering Tools* make use of the data inside the CDO to graphicsize the CDO. This is another important advantage of the current design.

3 Details of Classes in the Drawing Subsystem

3.1 Scene, Canvas and Postings

The design of the Drawing subsystem is centered around the *Scene.* There can be several Scenes in a *Job* - the list of Scenes is maintained bv the Scene *Tool.*

Each Scene is displayed in a *Viewport* - this is a one-to-one relationship. The *Viewport* and *Annotation* classes are derived from the *Posting class. Annotations* are used for displaying text only.

All *Postings* are posted on a Canvas. The *Canvas* has its own 2-D coordinate system and also maintains the list of *Postings* on it. Positions of Postings on the *Canvas* are specified using the canvas-coordinates. Each *Posting* has its own local coordinate system. The 3-D coordinate system of the Scene displayed in the *Viewport is* mapped into the 2-D local coordinates of the *Viewport.* A Canvas Viewing *Window* object provides a visible window onto the *Canvas,* The size and location of the *Canvas Viewing Window* determines the visibility of *Postings.* The *Postings* maybe moved and resized on the Canvas.

3.2 Drawable, Display Object List and DOLItems

The *DOLIt em* class is a convenient way of grouping together all the information required to render a CDO. Every CDO is derived also from the *Drawable* class in Drawing subsystem. All *Drawables* support a common interface for rendering purposes. In the Drawing subsystem, a CDO is identified as a *Drawable.*

The *GraphicsAttributes* class is the base class of a number of CDO-specific *GraphicsAttributes classes.* The CDO-specific *GraphicsAttributes class* contains information about graphics attributes such as colors and fonts to be used in the rendering. The *Graph icsMode* class keeps track of rendering modes and shading algorithms.

The *DOLItem* class contains as data a pointer to a *Drawable, a GraphicsAttributes* (CDO-specific), a *GraphicsMode* and several flags which indicate the rendering status of the object. .

The Scene maintains a list of *DOLItems.* It also has a *SceneGraphicsAttribute*, derived from all the CDO-specific *GraphicsAttributes*, and containing all the graphics attribute information associated with a Scene. This allows the Scene to set defaults to all the graphics attribute values of individual *Drawables*, and provide scene-wide values for all of them. Unless a graphics attribute value for an individual *Drawable* is specifically set, the required value is taken from the Scene *GraphicsAttribut e* while graphicsizing.

3.3 OIVNodeTable and RenderingTools

The rendering of CDOs is done using the *OpenInvent or* library [Silicon Graphics Inc., 1996]. OpenInventor maintains its own database of displayable objects in a scene. In OpenInventor too, there are scenes and displayable objects. However, this information is organised in the form of a tree structure, called a scenegraph [Wernecke, 1994]. The components of this tree, known as nodes, store the information needed for rendering the OIV scene. Subtrees may correspond to different objects in a scene. The root node of that subtree can be used as a handle to the OpenInventor database on the object it represents.

On-screen rendering of a scene is usually done using a window created through OIV or any of the predefine viewer objects available with OIV.

As mentioned, the rendering of CDOs is done using *OpenInventor* . To facilitate the easy switching of graphics libraries and to preserve the independence of the main classes in the Drawing subsystem from t bird party libraries, knowledge about Open Inventor in the Drawing subsystem is isolated in the *Rendering Tools* and the *OIVNode Table* classes.

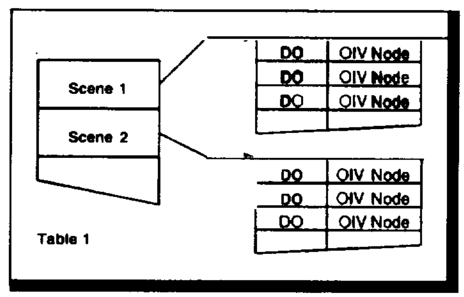


Figure 3: Table in *OIVNode Table class* for storing the OIV Nodes of *Drawables.*

The *OIVNode Table class* maintains a mapping between each *Drawable* and the root node of its corresponding OIV subtree. This is done on a per-Scene basis. Given a *Scene* and a *Drawable* in it, it is easy to find the corresponding OIV subtree rootnode from the table. This table can also be used to find the corresponding *Drawable* and Scene given an OIVNode - this is handy when objects are *picked* on the screen.

Scene 1	OIV Node	Viewer
Scene 2	OIV Node	Viewer
Scene 3	OIV Node	Viewer

Table 2

Figure 4: Table in *OIVNode Table* class for storing the Scene OIV Nodes and viewers

The *OIVNode Table class* also maps Scenes to their corresponding OIV rootnodes and OIV viewers. For every Scene there is an OIV root node as a handle to the corresponding OIV scenegraph, and an OIV viewer to display that particular scenegraph. This is stored as a separate table in *OIVNode Table.*

RenderingTools are the other classes in the Drawing subsystem that know about the graphics library. *RenderingTools* are derived, on a per-CDO basis from a common base class and therefore support a common interface. They are the classes which actually prepare the corresponding database in the graphics library for the *Drawables*, and therefore need to know how the OIV subtree for each *Drawable is* organised and how to populate it given a *Drawable.* Each *Rendering Tool is a* 'friend' C++ class of its corresponding CDO and is responsible for graphicizing instances of that type of CDO. They also have detailed knowledge of *GraphicsAttributes* and admissible *GraphicsModes* for each Drawable-derived class. All CDO-specific *Rendering Tools* exist on a one-per- *Job* basis. They are designed and implemented as *singleton* classes [Gamma et al., 1995].

4 Flow of Control During the Rendering of Scenes in a Job

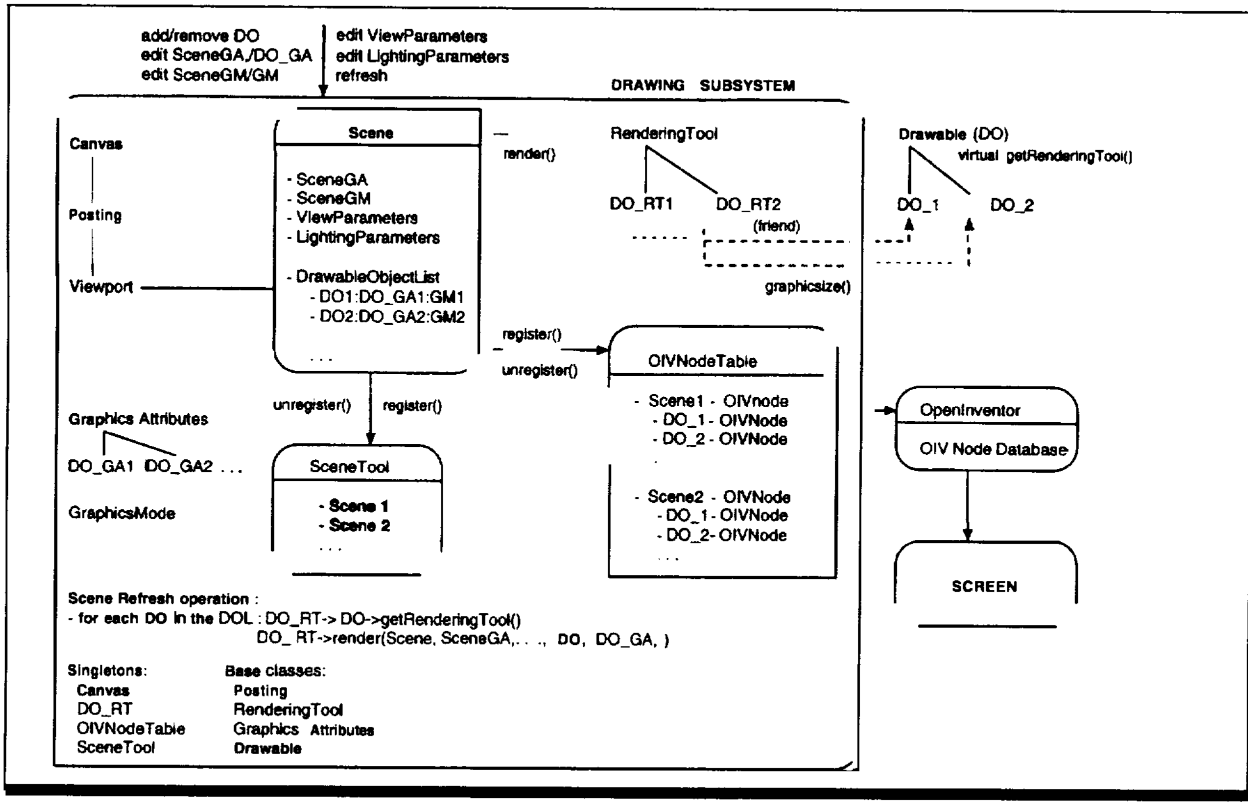


Figure 5: Interactions and Flow of Control in Drawing Subsystem.

Though there can be several Scenes per Job, there is always a current Scene rendering into the current *Viewport* of the current *Job*. However, when all Scenes in a Job are to be rendered, the list of Scenes maintained by the Scene *Tool is* traversed and each Scene in the list is asked to refresh itself.

Each Scene responds to this request by going through its DOL and asking the corresponding *Rendering-Tool* of each DOLItem's *Drawable* to populate the OIV database for that *Drawable*. The *Scene's* OIV node and the *Drawable*'s OIV Node are obtained from the *OIVNodeTable*; *so* is the OIV viewer into which the scene has to be rendered.

The *RenderingTool* knows how to graphicsize the *Drawable* appropriately, and once the DOL is fully traversed, all items are in the OIV database. For efficiency purposes there are steps to avoid unnecessary graphicsization, ensure the proper inclusion of all graphics attributes and graphics modes, and remove OIV nodes no longer needed. Once every *DOLItem* that needs to be graphicsized in the DOL of a Scene has been graphicized, the OIV viewer is asked to refresh the display. This completes the graphicsization of a *Scene*.

Once all Scenes are rendered into their appropriate *Viewports*, the rendering of the *Job* is complete.

5 Handling of Graphics Attributes

As mentioned, there are CDO-specific *GraphicsAttributes classes*, as *well* as a *Scene GraphicsAttribute* class which is derived from all of them. When a *Drawable* is introduced into the DOL, its own *GraphicsAttributes* are unset. The philosophy here is that the Scene *GraphicsAttribute* of the Scene should contain a valid default value for every *GraphicsAttribute* value possible. That default value is used for rendering till any of the *Drawable's GraphicsAttribute* values are explicitly set. In a similar vein, when a scene-wide value needs to be imposed for any graphics attribute, along with setting that value in the *SceneGraphicsAttribute* object, the corresponding values in the *GraphicsAttributes* of the *Drawables*' are unset, so that they are not used during the execution of next refresh command.

6 Conclusions

A Drawing subsystem has been implemented as per the design outlined in this paper. The use of OOD has been of great help in adding new classes into the subsystem or in making involved changes to one class modifying only the minimum number of other classes. Adding facilities for the rendering of a new CDO is a relatively easy matter in this design.

The main objective has been to keep the Drawing Subsystem independent of the underlying geometric modeler and the rendering library. Those are areas in which advances are made often, and if CoMeT's drawing subsystem has to keep up with the most advanced of these libraries, the changes required to convert from one to another have to be minimal. It may be concluded that this goal has been achieved in the design of the Drawing Subsystem.

References

[Gamma et al., 1995] Gamma, Erich., Helm, Richard., Johnson, Ralph., Vlissides, John., (1995) *Design Patt ems - Elements of Reusable Object-Oriented Software,* Reading, MA: Addison-Wesley.

[Headington et al., 1994] Headington, M. and Riley, D., (1994) *Data Abstraction* and *Structures Using C++*, Lexington, MA: D.C. Heath and Company.

[Lippman, 1993] Lippman, Stanley B., (1993) *C++* Primer, Second Edn., Reading, MA: Addison-Wesley.

[Silicon Graphics Inc., 1996] Silicon Graphics Inc., (1996) *OpenInventor.*, MountainView, CA.

[Somerville, 1996] Somerville, I., (1966) *Software Engineering,* fifth edn. Reading, MA: Addison-Wesley.

[Spatial Technology Inc., 1996] Spatial Technology Inc., (1996) *A CIS.*, Boulder, CO.

[Wernecke, 1994] Wernecke, Josie., (1994) *The Inventor Mentor*, Reading, MA: Addison-Wesley.

[Wirfs-Brock et al., 1990] Wirfs-Brock, Rebecca., Wilkerson, Brian., Weiner, Lauren., (1990) *Design of Object-Oriented* Software, Englewood Cliffs, NJ: Prentice-Hall.

NASA ...

Barium Nitrate Raman Laser Development for Remote Sensing of Ozone

Christopher L. McCray and Thomas H. Chyba

Research Center for Optical Physics,
Department of Physics, Hampton University, Hampton, VA 23668
(757) 727-5922, (757) 727-5955 (FAX), T.H.Chyba@larc.nasa.gov

Introduction

In order to understand the impact of anthropogenic emissions upon the earth's environment, scientists require remote sensing techniques which are capable of providing range-resolved measurements of clouds, aerosols, and the concentrations of several chemical constituents of the atmosphere. The differential absorption lidar (DIAL) technique is a very promising method to measure concentration profiles of chemical species such as ozone and water vapor as well as detect the presence of aerosols and clouds[1]. If a suitable DIAL system could be deployed in space, it would provide a global data set of tremendous value. Such systems, however, need to be compact, reliable, and very efficient.

In order to measure atmospheric gases with the DIAL technique, the laser transmitter must generate suitable on-line and off-line wavelength pulse pairs. The on-line pulse is resonant with an absorption feature of the species of interest. The off-line pulse is tuned so that it encounters significantly less absorption. The relative backscattered power for the two pulses enables the range-resolved concentration to be computed.

Preliminary experiments at NASA LaRC suggested that the solid state Raman shifting material, $Ba(NO_3)_2$, could be utilized to produce these pulse pairs. A Raman oscillator pumped at 532 nm by a frequency-doubled Nd:YAG laser can create first Stokes laser output at 563 nm and second Stokes output at 599 nm. With frequency doublers, UV output at 281 nm and 299 nm can be subsequently obtained. This all-solid state system has the potential to be very efficient, compact, and reliable.

Raman shifting in $Ba(NO_3)_2$, has previously been performed in both the visible and the infrared. The first Raman oscillator in the visible region was investigated in 1986 with the configurations of plane-plane and unstable telescopic resonators.[3] However, most of the recent research has focused on the development of infrared sources for eye-safe lidar applications.[4,5]

Current Experiments

We have recently completed an initial set of measurements with the Raman oscillator in the visible region.[6] The experimental apparatus for these tests is shown in figure 1. A compact, Q-switched frequency-doubled Nd:YAG laser serves as a pump laser for the Raman oscillator. Its PRF can be set at 30/N Hz by triggering

the Q-switch on every Nth flashlamp pulse. A half-wave plate (HWP) and polarizer cube (P) combination provides variable attenuation for the pump beam. The two lens-telescope provides control over the pump beam collimation and diameter. The Raman oscillator consists of a high reflector (HR), an output coupler (OC) and the Raman medium. The $Ba(NO_3)_2$ crystal (1x1x5 cm) is mounted in a sealed housing filled with index-matching fluid. The exit windows are A-R coated. The HR is a mirror which is 99.5% reflective at the Stokes wavelengths, but 90% transmissive for the 532 nm pump beam. Three sets of output couplers were designed to optimize the generation of first Stokes and second Stokes wavelengths separately and both Stokes wavelengths simultaneously. Doubling crystals convert the visible wavelengths into the UV.

Figure 2 is a plot of output energy from the Raman oscillator at the second Stokes wavelength vs. 532-nm pump energy for two pulse repetition rates. Pump energy is measured inside the resonator just before the barium nitrate crystal. There is virtually no first Stokes output for this cavity mirror configuration. The beam diameter at the crystal is 2.5 mm. Pump fluence must be maintained below the damage threshold for this material, which is approximately $8 J/cm^2$.

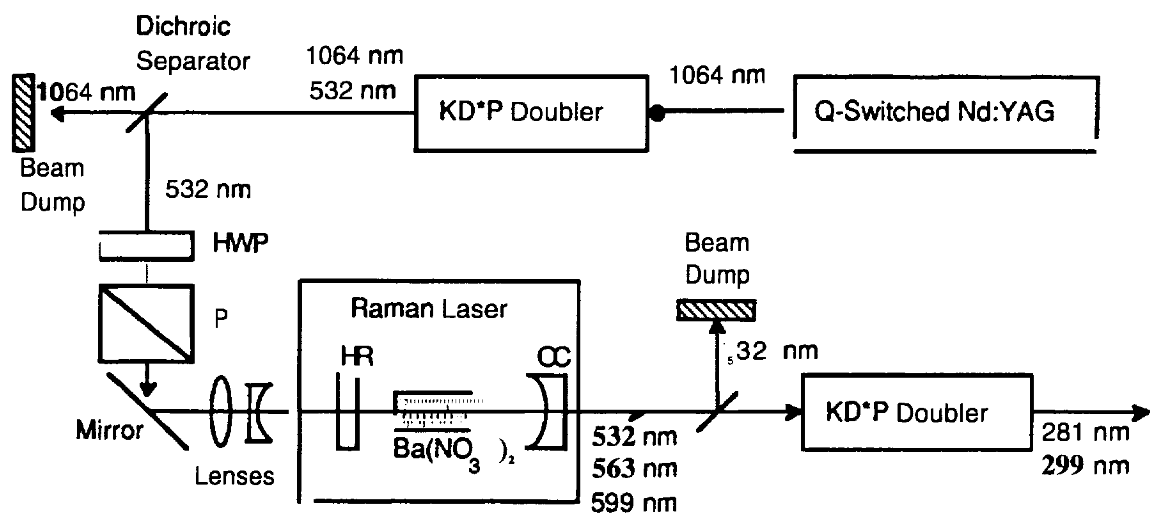


Figure 1. The barium nitrate Raman oscillator, configured for uv generation.

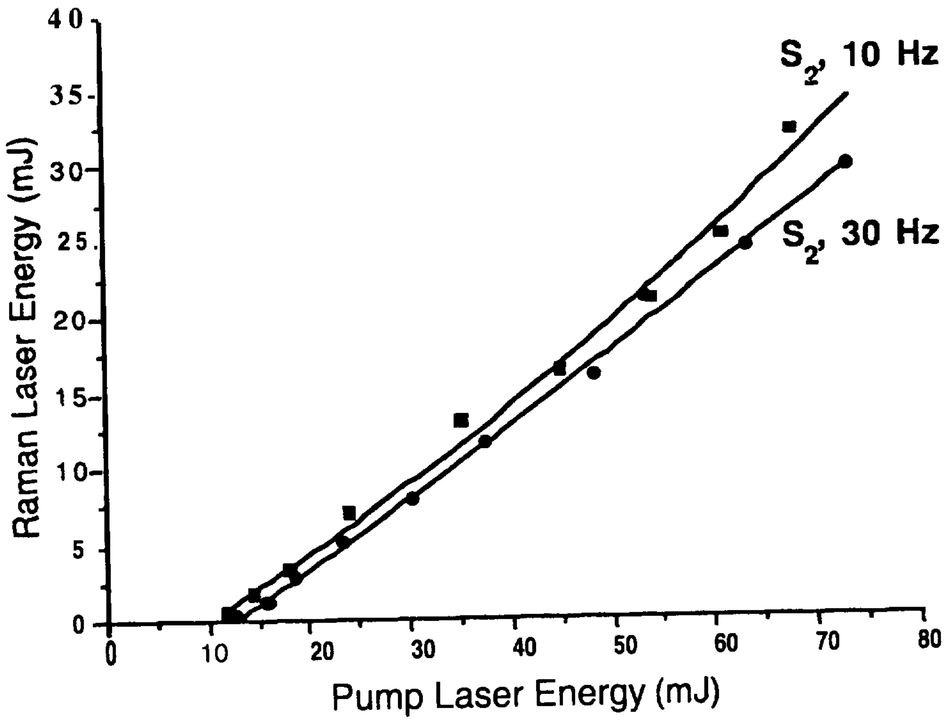


Figure 2. Output energy at the second Stokes wavelength vs. pump laser energy for two pulse repetition frequencies.

Conclusion

We are developing a compact, all-solid state UV source for DIAL measurements of ozone. We have Raman-shifted 532 nm pulses with up to 40% efficiency for the first Stokes and 48% efficiency for the second Stokes with our initial oscillator configuration. Beam divergence in the visible was measured to be 1.5 to 3.5 mrad with this particular pumping geometry.

At the conference, we will present results from our improved oscillator configurations for the visible and the UV.

Acknowledgements

The authors acknowledge Dr. Chuan He who performed the initial experiments on the Raman oscillator in our lab and Dale Richter, Science Applications International Corp., for many technical discussions. This research has been supported by NASA grant NAGU-2929.

References

1. see the review article by W. B. Grant, chapter 7 in Tunable Laser Applications, F. J. Duarte, cd., New York: Marcel Dekker, Inc., 1995, pp. 213-306.
2. D. Richter, W. Marsh, and J. Barnes, NASA Langley Research Center, private communication, 1995.
3. S. N. Karpukhin and A. I. Stepanov, Sov. 1. Quantum Electron. 16, 1027-1031, 1986.
4. J. T. Murray, R.C. Powell, N. Peyghambarian, D. Smith, W. Austin and R.A. Stolzenberger, Out. Lett, 20,1017-1019,1995.
5. P. G. Zverev and T. T. Basiev, Proceedings on Advanced Solid State Lasers, 24, 288-295,1995.
6. C. He and T. H. Chyba, Opt. Commun., in press, 1996.

URC97086

/ / / / -//
-/3' /43

On the Implementation of a Land Cover Classification System for SAR Images using Khoros•

Edwin J. Medina Rivera and Ramón Vasquez Espinosa

Abstract

The Synthetic Aperture Radar (SAR) sensor is widely used to record data about the ground under all atmospheric conditions. The SAR acquired images have very good resolution which necessitates the development of a classification system that process the SAR images to extract useful information for different applications. In this work, a complete system for the land cover classification was designed and programmed using the Khoros, a data flow visual language environment, taking full advantages of the polymorphic data services that it provides. Image analysis was applied to SAR images to improve and automate the processes of recognition and classification of the different regions like mountains and lakes. Both unsupervised and supervised classification utilities were used. The unsupervised classification routines included the use of several Classification/Clustering algorithms like the K-means, ISO2, Weighted Minimum Distance, and the Localized Receptive Field (LRF) training/classifier. Different texture analysis approaches such as Invariant Moments, Fractal Dimension and Second Order statistics were implemented for supervised classification of the images. The results and conclusions for SAR image classification using the various unsupervised and supervised procedures are presented based on their accuracy and performance,

1. Introduction

Traditionally, the source of region distribution information has been maps created by ground survey, An alternative approach could be the use of remote sensing and image processing techniques. In numerous studies such remotely sensed data has been used to accurately map vegetation, crop and other land-cover types. Most of these studies have been performed using LANDSAT, SPOT and AVHRR data, and in other cases using SAR data.

But, visible spectrum images of the ground depend on the atmospheric conditions of the area under analysis. Then, important information about the ground can not be taken all the time. A possible solution is the use of an active microwave sensor such as the Synthetic Aperture Radar (SAR).

In this paper, a land cover classification system designed for SAR images is presented. This system is built under the Khoros environment. The classification system was designed, and analyzed based on supervised and unsupervised classification approaches. The knowledge obtained with this work could highlight fundamental computational issues like memory size and CPU execution time. In addition, important parameters related to the implementation structure of some algorithms can be analyzed and possible relations can be established between them and the response of the land cover classification system.

2. Data Acquisition

The data used in this project was obtained from the Jet Propulsion Laboratory homepage. The image was created using data from the Spaceborne Imaging Radar C/X-Band Synthetic Aperture Radar (SIR-C/X-SAR) [1]. The image is a false-color composite of the Mammoth Mountain area in California's Sierra Nevada Mountains centered at 37.6 degrees north, 119.0 degrees west, (see Fig 1). It was acquired onboard space shuttle Endeavour on its 67th orbit on

---

•This work is partially sponsored by NASA Grant NCCW -0088.

April 13, 1994. In the image, red is C-band HV-polarization, green is C- band HH-polarization and blue is the ratio of C-band VV- polarization to C-band HV-polarization.


Figure 1: Original SAR image.

3. Supervised Classification System

Image classification systems are basically divided into two categories: supervised and unsupervised classification systems [2]. In the supervised systems, there is a supervisor to teach the system first, how to classify known set of images, and then the system goes ahead freely classifying other images, Usually it needs a *priori* information derived from field survey, photo interpretation and other sources about regions of the image. The supervised classification system is composed of two stages: the training and the classification stages. In the training stage, the user teaches the system how to classify the different regions of the image. This process is based on the proper selection of features and the fine tuning of several parameters of the system. The main function of the training stage is the selection of the objects which will determine the composition of the classes. First, the input SAR image is partitioned using a simple segmentation algorithm called labeling [3], The labeling algorithm performs segmentation based on the 4 (or 8) connectivity of a given pixel with its neighbors. Second, each region of the segmented image is analyzed to extract features using a texture analysis method.

The texture analysis was performed using invariant moment, fractal dimension and Gray level cooccurrence matrix (GLCM) texture features based methods. Lets take an overview of the different techniques.

3.1 Invariant Moments

The two dimensional (p+q)th order moments are defined as

$$m_{pq} = \sum_i \sum_j i^p j^q I(i,j) \quad \text{p, q} = 0,\ 1, 2 \ldots \tag{1}$$

where I(i,j) is the image of the object. The center of gravity or mean can be calculated from

$$\bar{x} = \frac{m_{10}}{m_{00}} \qquad \bar{y} = \frac{m_{01}}{m_{00}}. \tag{2}$$

The central moments are defined as

$$\mu_{pq} \cdot \sum_i \sum_j (i-\bar{x})^p \;\_\; \bar{y})^q I(i,\ j) \tag{3}$$

which are invariant under translation,

Similarly, a more general set of invariant moments can be obtained. These are the seven low-order central invariant moments Ml , ... M_7[4] and they are independent of size, orientation and position. The first six

moments are invariant under rotation and reflection, However, the last one, M_7 , is sensitive to reflection, its value changes sign for a reflected image of the object, but its magnitude remains unaltered.

3.2 Fractals

One important parameter used to characterize a surface is the fractal dimension. The concept of fractals is based on the continuous repetition of a mathematical pattern on a given random selected location of a space, Classical geometry is based on integer dimensions; fractal geometry, instead, deals with non-integer dimensions. That is, while a line has one dimension and a plane has two dimensions, a fractal curve will have a dimension between one and two depending on the intricacy of it, It has been proved that nature exhibits some structures similar to fractal objects and fractal models are used to synthesize and analyze different textures with good results [5], It has been proved that there is a correlation between the roughness of the surface and the value of the fractal dimension, As the fractal dimension increases more complex is the fractal and in the same way the roughness of the texture. The fractal dimension can be obtained using the following relation,

$$D = \frac{\ln(N_r)}{\ln\frac{1}{(r)}} \tag{4}$$

where r is the reduction ratio, N_r is the number of copies and D is the fractal dimension.

The estimation of the fractal dimension of an image is based on the probability that there are m pixels within a window of size L centered on a pixel from a particular class. For a selected range of window sizes (L), the window is centered on the first occurrence of the pixel belonging to a particular class. The number of pixels within a window of size L, belonging to a specific class are counted (including the center pixel of the window), and a histogram is formed as the window is moved over the image. This histogram produced as part of the method represents the total number of occurrences, m, of a class of pixels in a window of size L.

3.3 Gray-level cooccurrence matrix

The second order statistics provide a simple approach to capture the spatial relationship in a texture pattern. The GLCM considers the distribution of the intensities as well as the positions of pixels with equal or nearly equal intensity values. The GLCM is calculated as the number of times a given pair of pixels separated by a distance d are found in the image according to its corresponding locations[6]. The GLCM can be computed independently for each one of the nearest neighbors based on the angles 0°, 45°, 90° or 135°. Once the cooccurrence matrix is generated the Haralick texture features can be measured. The complete list of Haralick's features can be seen in reference [7].

4. Unsupervised Classification System

The unsupervised classification system does not require a supervisor to teach the system. It is based on clustering algorithms which determine the composition of the classes. Initially, C points are selected to serve as candidate cluster centers. Each pixel is examined and they are assigned to the nearest candidate cluster. This assignment would be made on the basis of the Euclidean or City distance measure. A new set of means are computed from the previous grouping produced. If the new means converge with the previous ones the procedure terminates. Otherwise, the procedure returns to second step with the current mean set. The next step after the clustering process is the classification of the clusters in a particular class.

One of the most popular classification algorithms is the Weighted Minimum Distance. It can distinguish a single class from the rest of the data or multiple classes from each other. Each pixel is assigned to the nearest candidate class

based on the minimum Euclidean Distance or minimum Euclidean Squared Distance. Another unsupervised Khoros classification algorithm is based on the Localized Receptive Field (LRF) Training/ Classifier.

Localized Receptive Field is a two layer topology units which consists of a single layer of self organizing, " localized receptive field" units and a single layer of perception. The single layer of perception units use the Least Mean Square or the Adaline learning rule to adjust weights. The weights are trained and then similar images may be quickly classified based on the training data set.

5. Methodology

In the supervised classification, the SAR image was segmented in 28 different regions using the labeling algorithm based on 4-connectivity neighbors. Then, the texture analysis techniques were applied to the segmented image and a matrix of features was produced for each method. The feature under analysis had to be extracted from its corresponding matrix and a histogram of the feature through the different regions of the image was displayed, The features were the seven invariant moments, the fractal dimension, the second angular moment and contrast with horizontal neighbors. The histogram is analyzed and 4 objects were selected based on the statistical frequency of the features, see Table 1 and Table 2.

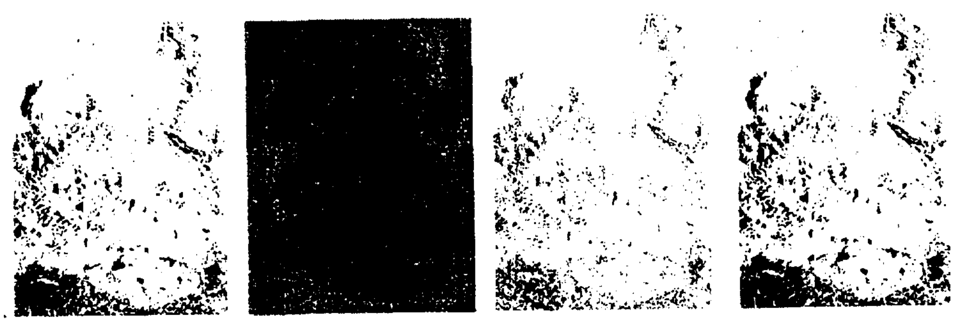
class	Object	Object Number	Feature Value
0	snow	5	0.2373
1	bare ground	7	0.2798
2	open water	21	05195
3	unknown	23	0.6672

Table 1: Prototyping table for Invariant Moments

Class	Object	Object Number	Feature Value
0	snow	11	1.5483
1	unknown	12	1.6508
2	bare ground	14	1.7666
3	open water	24	1.3765

Table 2: Prototyping table for Fractal Dimension

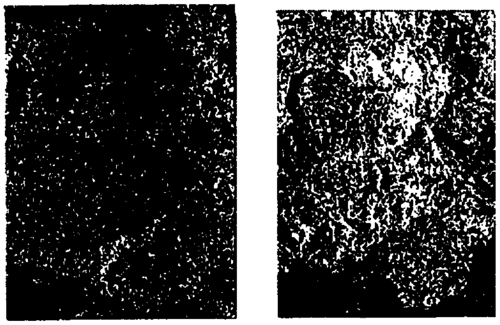
Features with higher frequencies are selected as the prototype classes or training data. After the training, classification stage follows the same steps as the training stage. The SAR image is segmented in 28 regions and then the texture analysis algorithm is applied for each region of the image. The vector produced in the training and the classification stages are compared using the minimum distance classifier and the different regions were classified, (see Fig. 2).



(a) (b) © (d)

Figure 2: Classified images using (a) invariant moments, (b) fractal dimension, (c) second angular moment and (d) contrast.

In the unsupervised classification, the K-means and the Isodata 2 algorithms were applied to the SAR image [81 [9]. Then, the cluster centers were labeled and combined with the original image to produce a multiband image. This image was used on the Weighted Minimum distance and the Localized Receptive Field classifiers, (see Fig. 3).



(a) (b)

Figure 3: Classified images using (a) K-means/ LRF classifier and (b) ISO2/Weighted minimum distance classifier.

Conclusions

The Invariant moments approach is the simplest one and could not correctly classify the open water region, It could discriminate between snow and bare ground very well, that is, it can discriminate between well defined regions.

Fractal dimension proved to be better than the invariant moment approach to classify a region. It produced good enough results in the classification of the open water region and the time to analyze the image was comparable with the invariant moments and less than the concurrence matrix method. The fractal dimension measures quantitatively the roughness of a surface in a way similar to human perception of the texture. Unfortunately, the roughness is not the only important texture feature, because there are quite different textures with the same fractal dimension. Fractals do not take into account the distribution of pixels; so its ability to discriminate different textures is not efficient. Another important issue to keep in mind is that the range in which we plot the log Nr vs log (l/r) is linear. Real objects has a range limited by lower and upper bounds.

The matrix obtained from the cooccurrence matrix approach is a square matrix with more elements than the original image. Note, that the number of operations required to process an image is directly proportional to the size of the original image. The computational dynamic storage cost of this process depends on keeping two consecutive lines of the image data in main memory, so storage constrains are determined only by the width of the image data.

References

[1] Jordan, R.L., Huneycutt, B.L. and Werner, M., " The SIR-C/X-SAR Synthetic Aperture Radar System ", *IEEE Transaction on Geoscience and Remote Sensing,* Vol. *33, No. 4,* Jul *95,* pp. *829-839.*

[2] Richards, J, A., *Remote Sensing Digital Image Analysis, An Introduction,* Springer-Verlag, USA, 1986.

[3] Gonzalez, R. C. and Woods, R. E., Digital Image Processing , Addison-Wesley Publishing Company, USA, 1993.

[4] Levine, M. D., Vision in Man and Machine, McGraw Hill, USA, 1980.

[5] Peitgen, H.O. and Saupe, D., The Science of Fractal Images, Springer-Verlag, USA, 1988.

[6] Khalaf, S., El-Galabi, M. and Abdelguerfi, M., "A parallel architecture for concurrence matrix computation ", *Proceedings of the 36th Midwest Symposium on Circuits and Systems, 1993,* vol. 2, pp. 945-948.

[7] Haralick, R.M., Shanmugam, K. and Dinstein I., " Textural Features for Image Classification ", *IEEE Transaction on system, man, and cybernetics, 1973,* Vol. SMC-3, No.6, Nov. 73, pp. 610-621.

[8] Genovese, G., Millet, M. and Schoenmakers, R.P.H.M. , " Segmentation and Interpretation of NOAA-AVHRR multitemporal NDVI based data sets in temperate regions ". *IGARSS '94. International Geoscience and Remote Sensing Symposium. Surface and Atmospheric Remote Sensing: Technologies,* Data *Analysis and Interpretation, 1994,* vol 4, pp. 2525-2527.

[9] Medina, E. J. and Vasquez, R., "Khoros classification algorithms applied to SAR images". *CRC 96. Computing Research Conference: Harnessing the information technologies,* April 1996, pp. 46-47.

NASA
S37/63

URC97087

Adaptive Fuzzy Control of a Direct Drive Motor *

E. Medina, Y. T. Kim, & M.-R. Akbarzadeh-T.
NASA Center for Autonomous Control Engineering
Electrical Engineering and Computer Engineering's Building
University of New Mexico, Albuquerque, NM 87131, USA
E-mail: emed@eece.unm.edu

Abstract

This paper presents a state feedback adaptive control method for position and velocity control of a direct drive motor. The proposed control scheme allows for integrating heuristic knowledge with mathematical knowledge of a system. It performs well even when mathematical model of the system is poorly understood. The controller consists of an adaptive fuzzy controller and a supervisory controller. The supervisory controller requires only knowledge of the upper bound and lower bound of the system parameters. The fuzzy controller is based on fuzzy basis functions and states of the system. The adaptation law is derived based on the Lyapunov function which ensures that the state of the system asymptotically approaches zero. The proposed controller is applied to a direct drive motor with payload and parameter uncertainty, and the effectiveness is verified by simulation results.

Keywords: Fuzzy Control, Adaptive Control, Stability, Direct Drive Motors.

Introduction

Direct drive motors have received increasing attention since they do not have a backlash or dead zone which are caused by gears. The high torque of the direct drive motor allows for the direct connection of the load to the motor axis. Since they are used in high-precision robot and machine tool applications, they must have high resistance to external disturbances. The absence of gear reduction leads to great sensitivity for the motor to variations in the load inertia. In fact, direct drive motors require more robust torque control than the conventional servo motors.

As a result, a linear controller cannot provide a good response under varying load conditions. Variable Structure System (VSS)-type self-tuning control [1], Bang-Bang control [2], and adaptive control [3, 4] have been proposed to handle such problems. However, none of the above mentioned approaches have taken advantage of the robustness of fuzzy logic in the controller design of direct drive motors.

This paper presents a state feedback fuzzy adaptive control method for position and velocity control of a direct drive motor for more robustness to system disturbances, The proposed control scheme does not require an accurate mathematical model of the system. It allows for integrating heuristic knowledge with mathematical knowledge. It performs well even when mathematical model of the system is poorly understood. The controller consists of an adaptive fuzzy controller and a supervisory controller. The supervisory controller requires only knowledge of the upper bound and lower bound of the system parameters. The fuzzy controller is based on fuzzy basis functions and states of the system. The adaptation law is derived based on the Lyapunov function which ensures that the state of the system asymptotically approaches zero. The proposed controller is applied to a direct drive motor with payload and parameter uncertainty, and the effectiveness is verified by simulation results.

Our objective is to control a Direct Drive motor (DD) to follow a desired trajectory. The high torque of DD motors allows for the direct connection of the load to the motor axis. Because of this, the motor becomes very sensitive to the load inertia applied. Therefore, we must design a controller that is robust with respect to the applied loads to the motor. In this paper a first-type adaptive fuzzy controller,i.e. one in which the adaptive parameters are linear, is used to meet our objective.

In the following section a linear model of a direct drive motor is presented. The next section shows the derivation of the first-type fuzzy adaptive controller, where stability is guaranteed by the Lyapunov based

*This work was supported in part by NASA contract # NCCW-0087

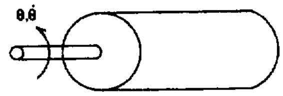


Figure 1: Direct-drive motor and the control parameters

supervisory controller. Then simulations of the DD motor which shows robustness of the fuzzy controller to parameter variations within the model of the DD motor is shown.

Mathematical Model

A DD motor is modeled and a state-space model is developed. Figure 1 shows the control parameters, θ and $\dot{\theta}$, which correspond to $\dot{x}_1$ and $\dot{x}_2$, respectively. The mathematical model of a DD motor is represented by the following differential equations,

$$J\ddot{\theta} + D\dot{\theta} = \tau \tag{1}$$

where,

J - Inertia moment of the system load and rotor
D- Coefficient of viscous friction term
19- Angular displacement of the motor (output)
τ - Output torque of the motor (control input, $u = \tau$)

Describing (1) in state space form we have a second-order system of the form,

$$\dot{x}_1 = x_2 \tag{2}$$
$$\dot{x}_2 = f(x, \dot{x}) + bu \tag{3}$$

where f is an unknown function, b is an unknown constant, and u and y are the input and output, respectively. After substitution, the above equation become,

$$\dot{x}_1 = x_2 \tag{4}$$
$$\dot{x}_2 = \frac{-D}{J}x_2 + \frac{1}{J}u \tag{5}$$

from which,

$$f(x_1, x_2) = \frac{-D}{J}x_2 \tag{6}$$
$$b = \frac{1}{J} \tag{7}$$

b is lower bounded and f is upper bounded and both exist.

Stable Direct Adaptive Fuzzy Controller

BACKGROUND

The adaptive fuzzy controller designed for the control objective follows the work of Wang [5], refer to Figure 2 for a block diagram of the control scheme. The resulting control law is the summation of a basic fuzzy controller, u_c, and a supervisory controller, u_s,

$$u = u_c(\underline{x}|\underline{\theta}) + u_s(\underline{x}) \tag{8}$$

The basic fuzzy controller, $u_c(\underline{x}|\underline{\theta})$ is a fuzzy logic system of the form,

$$u_c(\underline{x}|\underline{\theta}) = \underline{\theta}^T \underline{\xi}(\underline{x}). \tag{9}$$

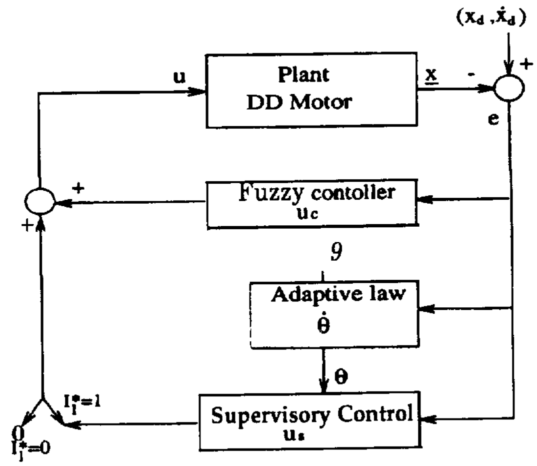


Figure 2: Block diagram of direct adaptive fuzzy control system

where $\underline{\theta}$ are the adjustable parameters, and the fuzzy basis function, $\underline{\xi}(\underline{x})$ is defined as,

$$\xi^{(l_1,l_2)}(\underline{x}) = \frac{\prod_{i=1}^{2}\mu_{F_i^{l_i}}(x_i)}{\sum_{l_1=1}^{m_1}\sum_{l_2=1}^{m_2}(\prod_{i=1}^{2}\mu_{F_i^{l_i}}(x_i))} \tag{lo}$$

where m_i is the number of fuzzy rule bases and μ_{F_i} are the fuzzy logic rules. By substituting (8) into (3), the state equation becomes,

$$x^{(2)} = \mathrm{f(a)} + b[u_c(\underline{x}|\underline{\theta}) + u_s(\underline{x})]. \tag{11}$$

If $f(\underline{x})$ and b are known, we know the control,

$$u^* = \frac{1}{b}[-f + y_m^{(n)} + \underline{k}^T\underline{e}] \tag{12}$$

will force the error to converge to zero, where $(\underline{e}) = (e, \dot{e})^T$ and $(\underline{k}) = (k_2, k_1)^T$. There fore, the error equation becomes, after adding and subtracting u^* to u,

$$\dot{\underline{e}} = \Lambda_c\underline{e} + b_c[u^* - u_c(\underline{x}|\underline{\theta}) - u_s(\underline{x})] \tag{13}$$

where Λ_c is a positive definite matrix and b_c is a vector defined respectively as

$$\mathrm{AC}= \begin{bmatrix} 0 & 1 \\ -k_2 & -k_1 \end{bmatrix} \tag{14}$$

$$b_c = \begin{bmatrix} 0 \\ b \end{bmatrix} \tag{15}$$

such that the values of k_1 and k_2 are chosen so the roots of $s^2 + k_1 s + k_2 = O$ are in the open left half plane.

Letting $V_e = \frac{1}{2}\underline{e}^T P\underline{e}$ it follows that,

$$\dot{V}_e = -\frac{1}{2}\underline{e}^T Q\underline{e} + \underline{e}^T P\underline{b}_c[u^* - u_c(\underline{x}|\underline{\theta})\ u_s(\underline{x})] \tag{16}$$

$$\leq -\frac{1}{2}\underline{e}\ \tilde{Q}\underline{e} + |\underline{e}^T P\underline{b}_c|(|u^*| + |u_c|) - \underline{e}^T P\underline{b}_c u_s. \tag{17}$$

The resulting equation for $\dot{V}_e$ is used to construct the supervisory control. We need for $\dot{V}_e \leq 0$ when $V_e > \bar{V}$, (a constant defined by the user). In order to meet the above objective, a supervisory controller, u_s, is designed according to the following assumptions. A function $f^u(\underline{x})$ and a constant b_L are determined such that $|f\ (\underline{x})| \leq f^U\ (\underline{x})$ and $0 \leq b_L \leq b$. *Now* u_s *is* constructed as follows:

$$u_s = I_1^* sgn(\underline{e}^T P\underline{b}_c)[|u_c| + \frac{1}{b_L}(f^U + |x_d^{(n)}| + |\underline{k}^T\underline{e}|)] \tag{18}$$

where $I_1^* = 1$ if $V_e > \bar{V}$, $I_1^* = O$ if $V_e \leq \bar{V}$.

Now the basic controller is replaced with a fuzzy logic system and an adaptive law is developed to adjust the parameter vector, $\underline{\theta}$. Next, the optimal parameter vector and the minimum approximation error are defined, respectively as,

$$\underline{\theta}^* \equiv argmin_{|\underline{\theta}| \leq M_\theta} [sup_{|\underline{x}| \leq M_x} |u_c(\underline{x}|\underline{\theta}) - u^*|] \tag{19}$$

$$w \equiv u_c(\underline{x}|\underline{\theta}^*) - u^* \tag{20}$$

where M_θ and M are constraints defined by the user. The error equation is rewritten as

$$\underline{\dot{e}} = \Lambda_c \underline{e} + b_c \underline{\phi}^T \underline{\xi}(\underline{x}) - b_c u_s - b_c w \tag{21}$$

where $\phi \equiv \underline{\theta}^* - \underline{\theta}$ and $\underline{\xi}(\underline{x})$ is the fuzzy basis function defined in (10). Since Λ_c is a stable matrix there exists a unique positive definite symmetric $2x2$ matrix P which satisfies the Lyapunov equation

$$\Lambda_c^T P + P\Lambda_c = -Q. \tag{22}$$

For a more detailed description on the derivation of the adaptive fuzzy controller and the adaptive law refer to [5].

DESIGN

The first-type direct fuzzy adaptive controller, is designed to control the direct drive motor. The controller is designed to be stable in the sense of Lyapunov, where the Lyapunov equation must be satisfied. The matrix P *was* arbitrarily chosen to satisfy the Lyapunov equation as,

$$P = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}. \tag{23}$$

The Lyapunov equation was solved using matlab with the P and Λ_c matrices as defined earlier. In our case, we chose $k_1 = 2$ and $k_2 = 1$ for the Λ_c matrix, which placed the roots of $s^2 + k_1 s + k_2 = O$ at s = −1, −1. The resulting matrix Q was

$$Q = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \tag{24}$$

a symmetric positive definite matrix, thereby satisfying the criterion for stability in the sense of Lyapunov.

The values of $f^U(\underline{x})$ and b_L are calculated. From the state equations we see the values of $f^U(\underline{x})$ and b_L are dependent on the J and D parameters of the motor. The parameters are motor dependent. In our case the values of J and D were chosen, respectively as $0.3 \leq J \leq 1.0$ and $0.001 \leq D \leq 0.1$. Therefore the values for $f^U(\underline{x})$ and b_L are $0.001x_2$ and 1, respectively.

Next, three fuzzy sets for xl and x_2 whose membership functions uniformly cover U were defined as

$$\begin{aligned} \mu_{F1}(x) &= 1/(1 + exp(x + 2)) \\ \mu_{F2}(x) &= exp(-x^2) \\ \mu_{F3}(x) &= 1/(1 + exp(-(x - 2))) \end{aligned} \tag{25}$$

Fuzzy basis functions are constructed using the above three fuzzy membership functions according to the relation (10).

The following adaptive law adjusts the state vector θ

$$\underline{\dot{\theta}} = \gamma \underline{e}^T \underline{p_n} \underline{\xi}(\underline{x}) \tag{26}$$

where γ is a constant, $\underline{e}^T$ is the desired trajectory minus the actual trajectory, and $\underline{p_n}$ *is* the last column of *P*.

Simulation

The adaptive fuzzy controller was simulated on Matlab using the ODE45 command to solve the differential equations. An M-file was written to describe the system of ordinary differential equations. A simulation was

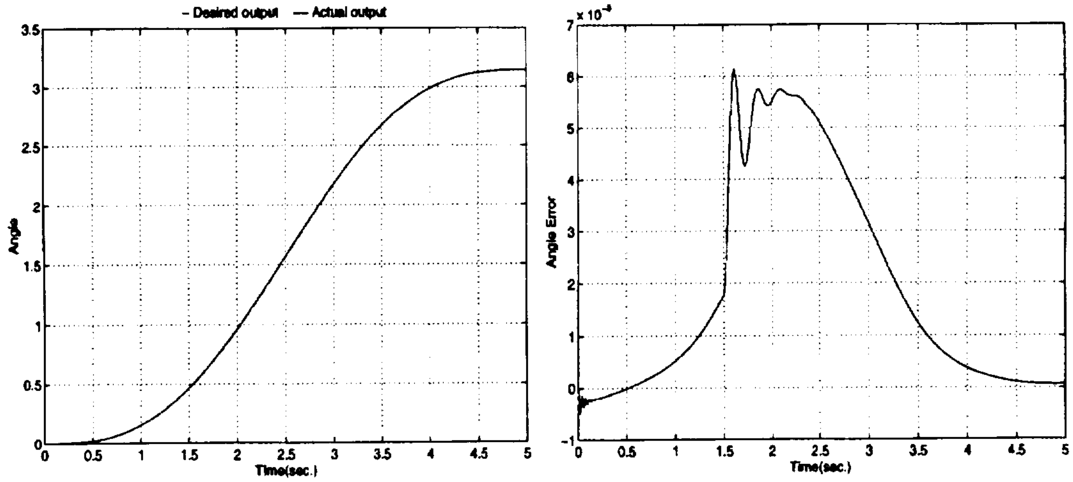


Figure 3: Plot of the angular and angular error response of the motor

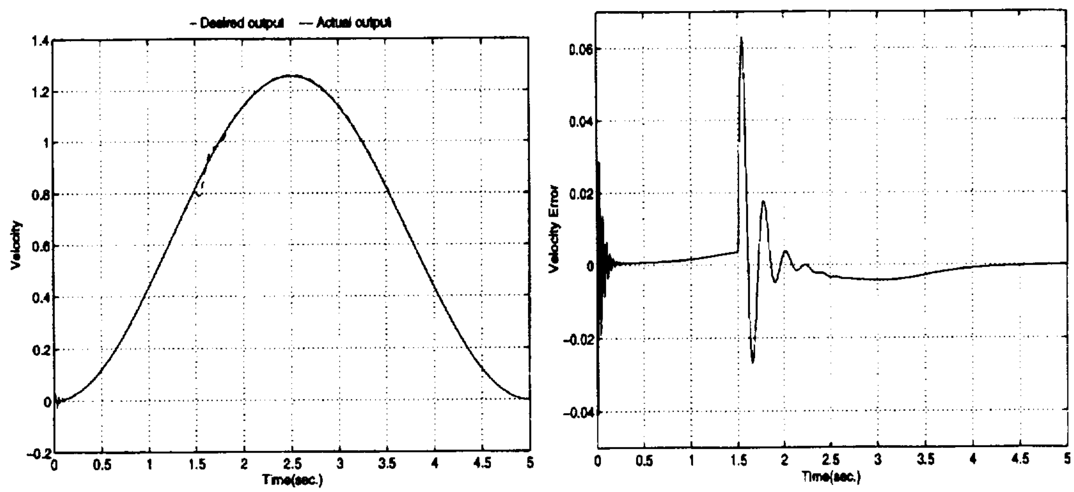


Figure 4: Plot of the velocity and velocity error response of the motor

run for 5 seconds using zero for the initial conditions of xl and x_2. A disturbance, a change in the load, was introduced after the simulation was started,

In order to simulate a change in load, the values of J and D were init ally 0.31 and 0.0077 then were changed 1.5 seconds after the simulation start ed to 0.5 and O. 1, respect ivel y. The desired trajectory is defined as

$$
\begin{aligned}
x_d &= 0.5(0.4\pi t - sin(0.4\pi t) \\
\dot{x}_d &= 0.2\pi(1.0 - cos(0.4\pi t))
\end{aligned}
$$

for x_1 and x_2, respectively.

Figure 3, shows the difference between the angle of the actual position and the desired position. It is shown that the difference in the desired and actual output is negligible. After the disturbance is applied, referring to Figure 5, the controller is able to adapt within 1 second.

In Figure 4, it is shown that the actual velocity response differs only at the start of the simulation and when the disturbance is introduced to the system. Figure 4 shows the adaptive fuzzy controller is robust with respect to variations in the motor load.

Figure 5, shows the adaptive fuzzy controller is robust with respect to variations in the motor load. Looking at the graph we are able to deduce the adaptive fuzzy controller is able to stabilize the system after the disturbance is introduced within 5 seconds. The controller quickly adapts to compensate for the changing load/inertia condition.

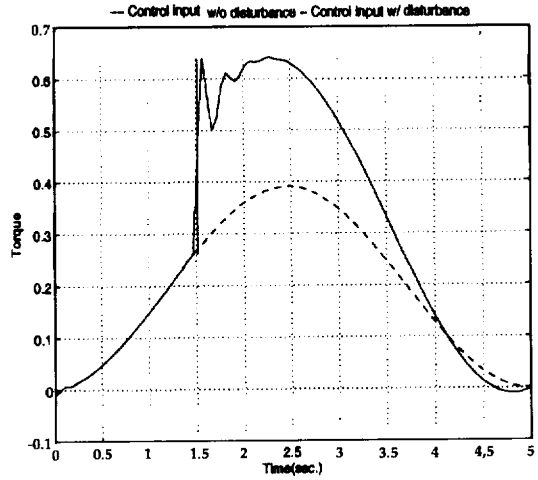


Figure 5: Plot of the control input DD with and without a disturbance used to control the DD motor

Conclusion

In most practical systems, some type of a mathematical model is available. Although this mathematical model is often corrupted by unknown parameters, disturbances, etc., it still constitutes a very important portion of a systems engineer's knowledge base. In this paper, heuristic knowledge (fuzzy logic rule base) is allowed to combine with a model-based controller to achieve a more robust control system.

A first-type fuzzy adaptive control method was introduced for position control of a direct drive motor. The control method presented does not require an accurate mathematical model of the control system, it is capable of incorporating fuzzy control rules directly into the controller and guarantee global stability of the resulting closed-loop system in the *sense* that all signals are asymptotically convergent. The controller was adaptively changed by monitoring the difference between the desired trajectory and the actual trajectory.

The control method discussed was shown to be robust with respect to the load on the motor, i.e. the system can be adaptively stabilized even though the load on the motor changes.

Future research topics include experimentally verifying the results of the simulation, using the DD motor and hardware available, to run real-time experiments. Also in future research, other autonomous methods will be implemented for generating the rule bases used to design the controller.

References

[1] K. Furuta, K. Kosuge, & K. Kobayashi, "VSS-Type Self-tuning Control of Direct Drive Motor," *Proceedings of the IECON*, pp. 281-286, 1989.

[2] K. Furuta & S. Kobayashi, 'Bang-Bang Position Control of Direct Drive Motor," Proceedings *of the IECON, pp. 148-153, 1990.*

[3] Y. T. Kim, B. C. Kim, & H. D. Lee, 'Adaptive Control of a Direct Drive Motor," *Proceedings of the KIEE*, Korea, *1994.*

[4] K. Furuta, K. Kosuge, & K. Kobayashi, 'Self-tuning, Sliding Mode and Adaptive Control for Direct Drive Motor," *Proceedings of the IECON*, pp. 459-465, 1988.

[5] L.-X. Wang, "Adaptive Fuzzy Systems and Control: Design and Stability Analysis," *Prentice Hall International Inc.* Englewoods Cliffs, NJ, 1994.

/ / / 209/62

URC97088

ENDOTHELIAL CELL MORPHOLOGY AND MIGRATION ARE ALTERED BY CHANGES IN GRAVITATIONAL FIELDS.

Caroline Melhado[§], Gary Sanford[§] and Sandra Harris-Hooker

Departments of Biochemistry[§] and Medicine[‡]
Morehouse School of Medicine
Atlanta, GA 30310.

INTRODUCTION: Many of the physiological changes of the cardiovascular system during space flight may originate from the dysfunction of basic biological mechanisms caused by microgravity (1 -4). The weightlessness affects the system when blood and other fluids move to the upper body causing the heart to enlarge to handle the increased blood flow to the upper extremities and decrease circulating volume. Increase arterial pressure triggers baroreceptors which signal the brain to adjust heart rate. Hemodynamic studies indicate that the microgravity-induced headward fluid redistribution results in various cardiovascular changes such as; alteration of vascular permeability resulting in lipid accumulation in the lumen of the vasculature and degeneration of the the vascular wall (5), capillary alteration with extensive endotheilial invagination (6). Achieving a true microgravity environment in ground based studies for prolonged periods is virtually impossible. The application of vector-averaged gravity to mammalian cells using horizontal clinostat produces alterations of cellular behavior similar to those observed in microgravity (7). Similarly, the low shear, horizontally rotating bioreactor (originally designed by NASA) also duplicates several properties of microgravity (8). Additionally, increasing gravity, i.e., hypergravity is easily achieved, Hypergravity has been found to increase the proliferation of several different cell lines (e.g., chick embryo fibroblasts) while decreasing cell motility (9) and slowing liver regeneration following partial hepatectomy (1 O). The effect of altered gravity on cells maybe similar to those of other physical forces, i.e. shear stress. Previous studies examing laminar flow and shear stress on endothelial cells found that the cells elongate, orient with the direction of flow, and reorganize their F-actin structure, with concomitant increase in cell stiffness (11). These studies suggest that alterations in the gravity environment will changethe behavior of most cells, including vascular cells. However, few studies have been directed at assessing the effect of altered gravitational field on vascular cell fiction and metabolism, Using image analysis we examined how bovine aortic endothelial cells altered their morphological characteristics and their response to a denudation injury when cells were subjected to simulated microgravity and hypergravity.

EXPERIMENTAL METHODS: Bovine aorta endothelial cells (BAEC) were obtained from the NIGMS/Coriell Cell Repository. Cells cultured on collagen coated flasks and collagen coated microcarier beads and maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum and 1X antibiotics.

Morphological Studies: Post confluent, Confluent and sparse BAEC were cultured in flask, confluent culture were then subjected to denudation injury, both were then centrifuged at 6G & 12G (HGrav) or clinostat rotated at 30 rpm (Mgrav). Phase contrast photomicrographs were taken at various times following injury and used to assess the morphology of cells as they moved into the denuded area. Cultures were also rinsed with Hanks Buffered Salt Solution and fixed with 10°/0 buffered formalin/2.5% glutaraldehyde for examination by scanning electron microscopy, Sparse cultures and cultures grown on Cytodex 3 microcarier beads in the horizontally rotating bioreactor were assessed similarly.

Migration and Motility Studies: Confluent, 1 -day and 4-days postconfluent (PC) cultures were subjected to HGrav (6G) and MGrav treatment, in 10% or 0.5% FBS, for 24,48, and 72 hr. At each time, an area of the culture was denuded to assess the migration of cells; cultures were kept under HGrav or MGrav following denudation. The migration of cells into the denuded area was monitored by video recording of random fields along the denuded area; cultures were recorded then returned to HGrav or MGrav conditions in the shortest time possible (5-10 rein). Similar studies were conducted using sparse cultures. We also assessed the influence of increasing hypergravity (6 and 12G) and clnostat rotation on BAEC morphology and migration. Control cultures were treated similarly, except they weremaintained under standard cell culture conditions and normal gravity, the clinostat control cultures were vertically rotated.

RESULTS AND DISCUSSION:

Endothelial Cell Migration: 1 -D post confluent and confluent BAEC maintained in DMEM containing 10% FBS did not

show any appreciable differences in total distance migrated under control conditions, HGrav or MGrav , The migration of BAEC was affected by changes in the gravity environment when cultured in low FBS (0.5%). The migration of confluent and 1 -day PC BAEC were retarded by MGrav treatment but stimulated by HGrav when compared to controls (fig. 1). The migration of 4-day PC BAEC was retarded by both Mgrav and HGrav, with HGrav producing the greatest decrease in total migration. Both 12G and 6G Hgrav resulted in a 30-50% retarded migration during the early response period between 2-6hr, between 16 and 48 hr this was slightly reversedby 5- 10% .Increasing Hgrav from 6G to 12G showed no significant difference in migratory response for 0.5%FBS cultures, however there wa a 20% difference observed in the 10% FBS cultures (fig. 2 &6). The 30 rpm horizontally rotated clinostat MGrav simulation resulted in a 30°/0 enhanced migration for the 0.5% FBS cultures tduring the early reponse period of 0-6hr , contrarily the 10'% FBS cultures had a 30% retarded migratory response. MGrav retarded the migratory response by 50% for both the O. 5°/0 and 10 '/0 FBS treated cultures (fig. 3&7)

Morphological Assessment: Confluent BAEC subjected to HGrav and examined by scanning electron microscopy had less surface area, fewer membrane-bound vesicles, smaller and more flattened nuclei, and membrane ruffling around the edges when compared to control cells. MGrav treated BAEC cultures were less elongated and had a more cobblestone appearance than controls. Examination of BAEC microcarrier cultures sujected to prolonged horizontal bioreactor rotation, by scanning electron microscopy showed a loss of surface vesicles changes in the appearance of these cells under MGrav compared to controls (fig. 5)

Changes in Endothelial Cell Area: The results of the image analysis of endothelial cell motility are shown in Fig. 4. The 4-day PC BAEC have a 50% smaller mean cell area than confluent cell; this is seen for cell in 10% or 0.5% FBS. Increasing the serum from O. 5% to 10% resulted in a three-fold increase in cell area for confluent cell; PC BAEC cell area also increased but only by 40%. Analysis of cell area alteration in response to denudation under HGrav conditions, 4-day PC BAEC are nonresponsive to serum-induced increase cell area (or spreading). There a dramatic decrease in responsiness to serum-induced increase in cell area for confluent BAEC; cell area incrcased only 75% compared to 300% for control cells. MGrav inhibited the responsiness of PC BAEC to serum-induced increase in cell area but resulted in an actual reversal of the serum effect with confluent BAEC. In the latter case, cells area decreased by 40% when serum levels increased from 0.5% to 10% (fig.5). The sparse cultures subjected to 24,48 hr Hgrav simulations showed a differences in cell shape (circularity and rectangularity) ,cells speared flatened and elongated (fig. 8)

ACKNOWLEDGEMENT: These studies were supported by grants NASA NAG 9-644 and NCCW 0085, and NIH/RCMIRR03034

REFERENCES

1. Rijken PJ, et al. (199 1). Aviat. Space Environ. Med. 62:32-36.
2. Kumei Y, et al. (1989). J. Cell Sci. 93:221-226.
3. Cogoli A, Tschogg A and Fuchs-Bislin P (1984). Science 225:228-230.
4. Gruener R and Hoeger G (1990). Am. J. Physiol. 258: C489-C494.
5. Doty S, E Holton, G Durnova and A Kaplansky. (1 990), FASEB J. 4:16.
6. Philpott D, I Popova, K Kate, et al. (1 990). FASEB J. 4:73.
7. Nixon J, R Murray, C Byrant, et al. (1979). J. Appl. Physiol. 46:541.
8. Nilsson J, Volk-Jovinge S, Svensson J, Landou C, DeFaire U and Hamsten A (1 992). J. Intern. Med. 232:397-404.
9. Rosen EM, Meromsky L, Setter E, Vinter DW and Goldberg ID (1 990). Exp. Cell Res. 186:22-31.
10. Kropacova K, et al. (1988), The Physiologists 31, S75-S76.
11. Girard PR and Nerem RM (1995). J. Cell Physiol. 163:179-193.

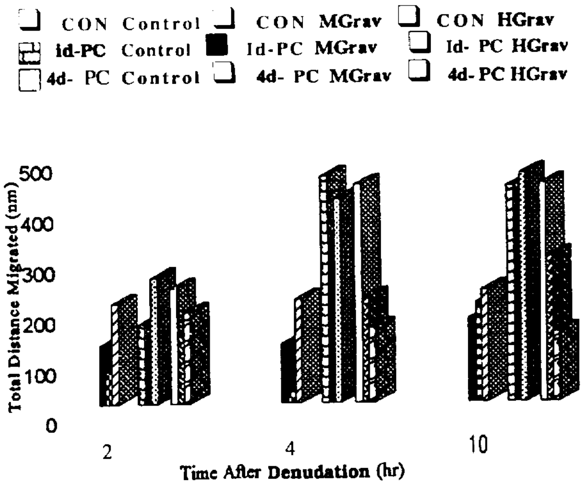


Figure 1. Alterations in BAEC migration induced by culture under HGrav and Mgrav conditons and 0.5% FBS: Dependency of BAEC migration on the degree of confluency.

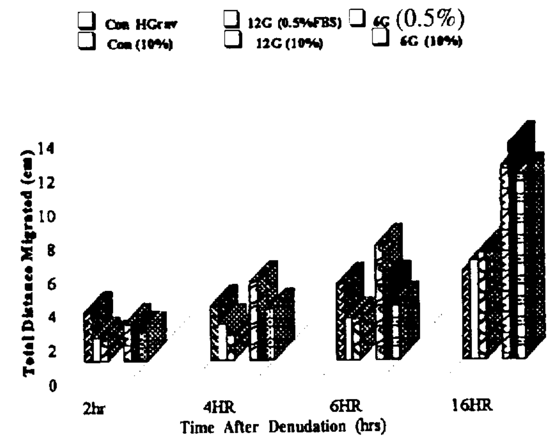


Figure 2. Alters in BAEC migration under 6 and 12G HGrav in 10% and 0.5% FBS.

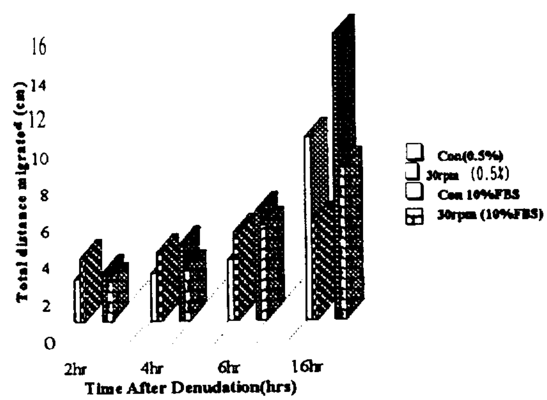


Figure 3. Alterations in BAEC migration under clinostat rotation in 10% and 0.5% FBS.

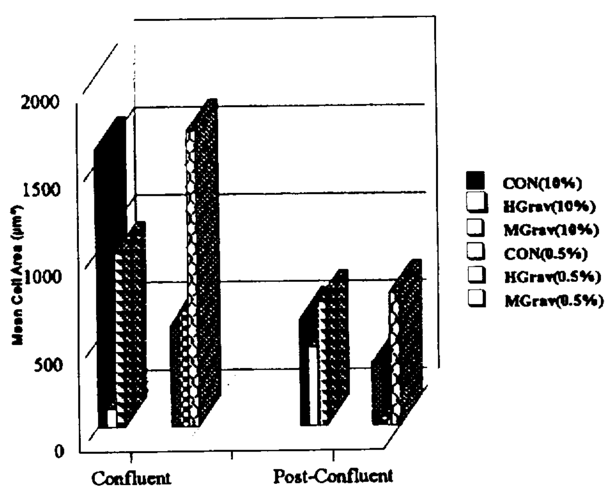


Figure 4. Changes in migrating BAEC area following culture under Hgrav and Mgrav conditions: dependency on the degree of confluence and serum levels.

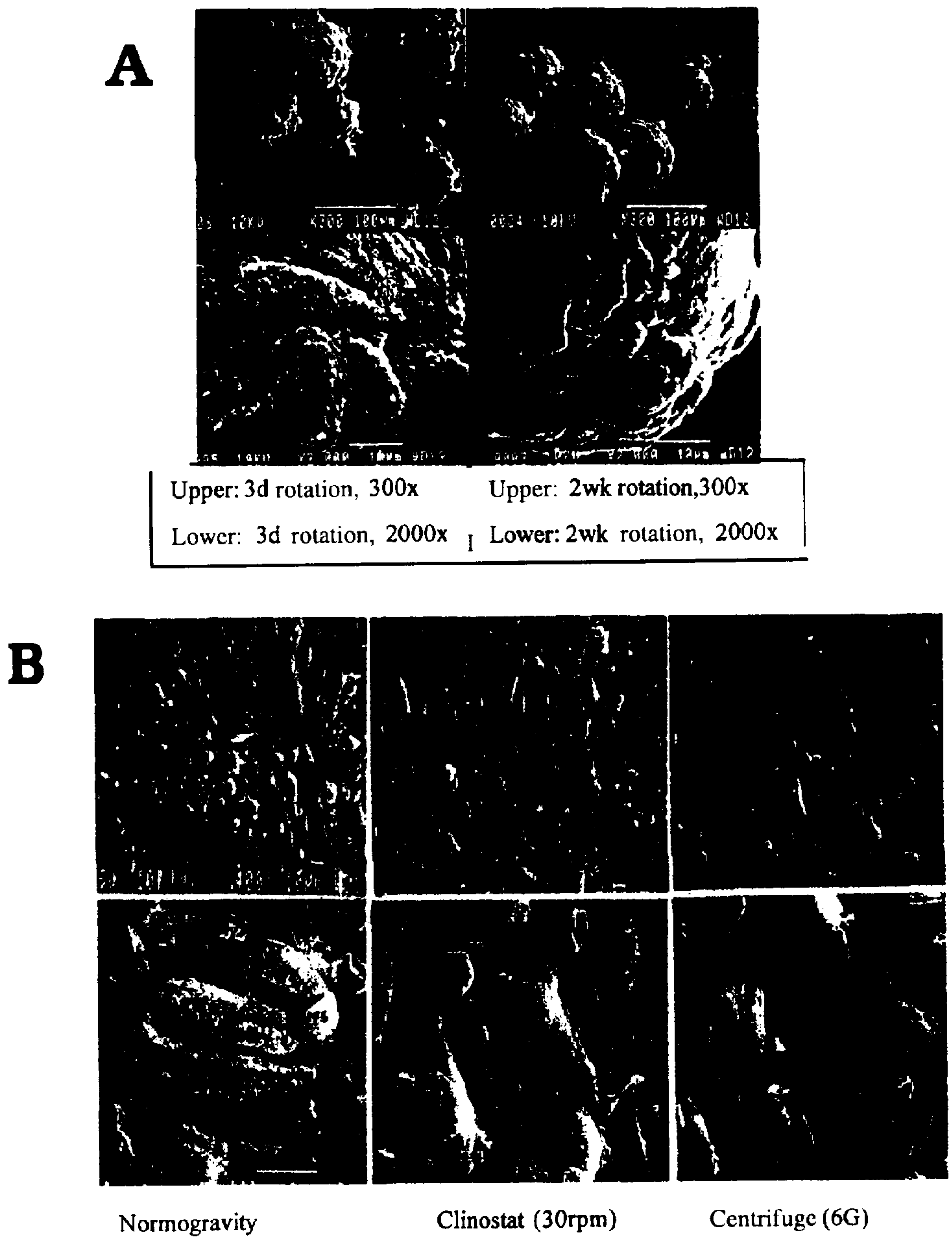
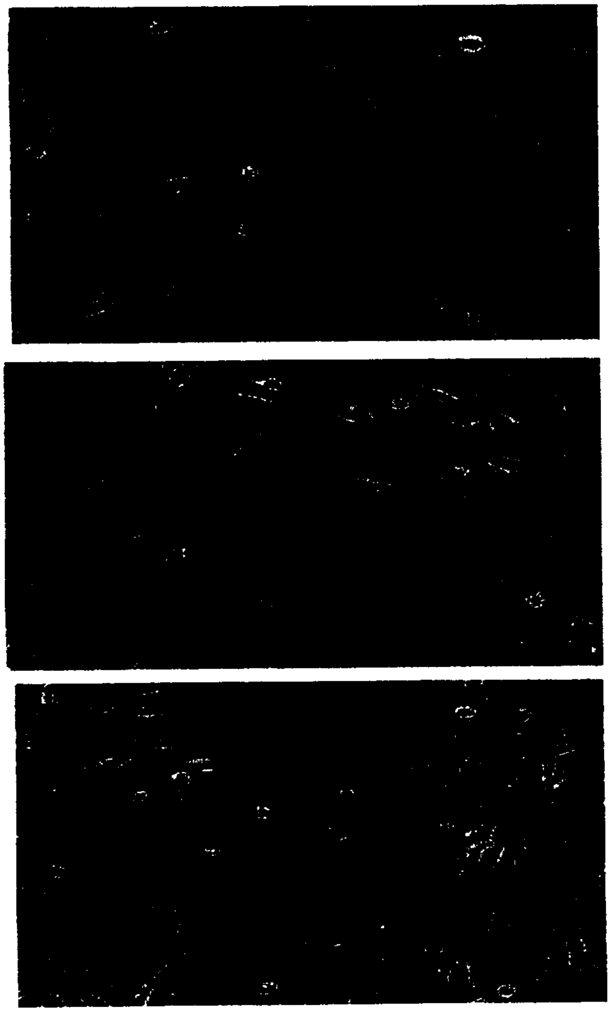


Figure 5. Scanning electron micrograph of BAEC cultured on Cytodex 3 microcarrier beads in the horizontally rotating bioreactor (A), and in flasks (B).



AE

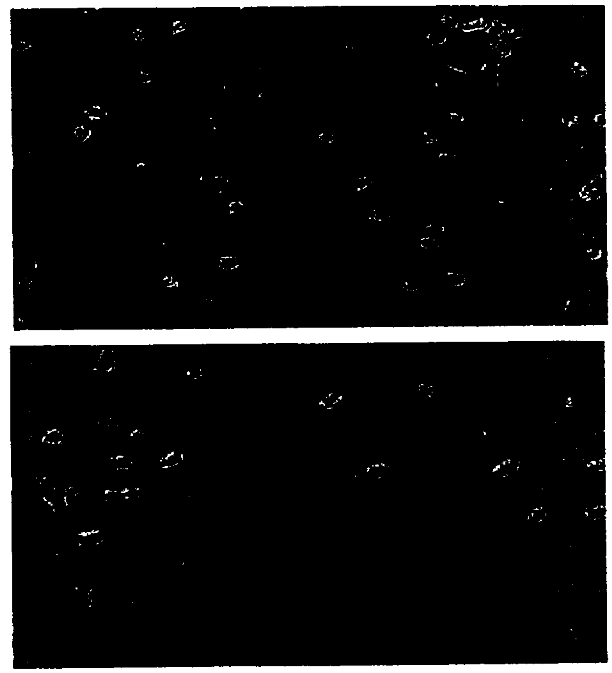


Figure 7. BAEC sheet migration after 16 hr normogravity (upper) and simulated microgravity (lower).

A

Figure 8. The change in BAEC morphology induced by 48 hr of hypergravity-6G (B) compared to non-centrifuged control (A). Sparse BAEC cultures were photographed under phase at a magnification of 250x.

Page intentionally left blank

URC097089

Modeling of the First Layers in the Fly's Eye

J. A. Moya[1], M. J. Wilcox[2] and G. W. Donohoe[1]

Depts. of Electrical and Computer Engineering and Psychology*

NASA ACE Center
Dept. of Electrical and Computer Engineering
University of New Mexico, Albuquerque, NM, 87131-1356.
jamoya@eece.unm.edu wilcox@cyber.unm.edu donohoe@eece.unm.edu

Abstract Increased autonomy of robots would yield significant advantages in the exploration of space. The shortfalls of computer vision can, however, pose significant limitations on a robot's potential. At the same time, simple insects which are largely hard-wired have effective visual systems. The understanding of insect vision systems thus may lead to improved approaches to visual tasks. A good starting point for the study of a vision system is its eye, In this paper, a model of the sensory portion of the fly's eye is presented. The effectiveness of the model is briefly addressed by a comparison of its performance to experimental data.

1 Introduction

One of the most important sources of information about our environment is certainly vision. Thus, it seems reasonable to endow robots with sight. However, computer vision has proven a difficult problem. At the same time, insects which are essentially hard-wired can solve many visual tasks. These include problems requiring pattern recognition, the tracking of objects and the selection of intercept courses [2,3,10]. Thus, the study of insect vision may lead to new approaches to visual problems.

A most important part of a vision system is its photoreceptor layer. This layer provides the sensory input and therefore sets limits on the performance of the system. Thus, a good starting point for the study of a vision system is in the most distal parts of the eye.

If one wishes to study an insect eye, the eye of the fly is an excellent choice. It is experimentally convenient and much literature on its eye has been produced. In this paper, after an introduction to the fly's eye, anew model of its photoreceptor layer is presented. The effectiveness of the model is also addressed.

2 The distal fly's eye

The eye of the fly is composed of a continuum of layers [1]. The most distal of these are the optical layer, the retina and the lamina ganglionaris, usually referred to as simply the lamina. The optical layer, the familiar compound structure seen at the eye's surface, is the most visible part of the eye. The term retina as applied to the fly differs from its use in vertebrates, referring only to the portions of photoreceptors distal to their own axons. The lamina contains the remaining portions of the photoreceptors and the second-order cells upon which the receptors synapse.

The structure of the fly's distal eye is depicted in fig. 1. The basis element called the ommatidium is composed of a pair of optical elements, the corneal facet and crystalline cone, and the retinal portions of the photoreceptors that lie immediately under this dioptric apparatus [1]. Each eye of a fly contains several thousand of these elements. The ommatidia are separated from one another by pigment cells which act as apertures and screen out stray light.

The optics are of high quality and deliver a well focused image to the tips of the photoreceptors within the ommatidium [6]. Further, because of their small size, chromatic aberrations are insignificant and depth of focus is relatively large [6,9].

As depicted in fig. 1B, the photoreceptors contained within the ommatidium, are arranged in an asymmetric trapezoid [1], Six peripheral receptors, referred to as R1 to R6, surround a pair of tandemly arranged central ones, R7 and R8. The central receptors and peripheral receptors are in general of different types and appear to serve different functions [5].

The fly's photoreceptor is composed of two parts, the rhabdomere and the cell body [1]. The rhabdomere acts as an absorbing waveguide [19] and transduces light into an ionic current [13]. The cell body can be further divided into three segments, the soma, the axon and the axon terminal [1]. The soma is the retinal part of the cell and receives the current generated in the rhabdomere. The terminal is the portion of the cell that transfers information onto second-order cells in the lamina. The axon is a conduit that connects the soma and terminal.

Unlike many animals with compound eyes, the fly has evolved an "open rhabdom" where every rhabdomere in the ommatidium is optically isolated from its neighbors [1]. In the fl y, these rhabdomeres also do not collect light

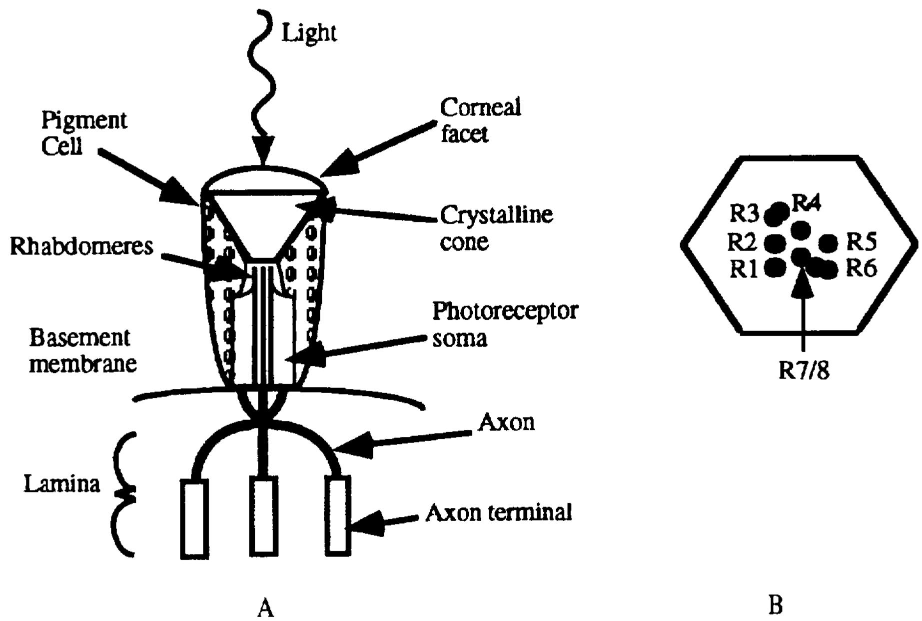


Figure 1 An ommatidium in the fly's eye and its axonal extensions into the next layer, the lamina. A. The ommatidium is composed of a set of photoreceptors, excluding their axons and axon terminals, and the optics immediately above them, the corneal lens and crystalline cone. Pigment cells form an aperture around this element and screen out stray light. The axonal portions of the receptors synapse with second-order cells in the lamina. **B.** The ommatidium as seen from the top of A.

from the same portion of space [8]. Although, as depicted in fig. 2, receptors from six, adjacent ommatidia do join together to sample similar regions. These receptors, however, do not sample concentrically [15]. The axons of these same photoreceptors, after penetrating the basement membrane, a high resistance barrier that separates the extracellular space (ECS) surrounding the ommatidia from the ECS of the lamina, also join together in the same element in the lamina [1].

In these lamina structures, new sets of RI -R6 photoreceptors are grouped. These grouped receptors are referred to as the neuro-ommatidium. The axon terminals in a single neuro-ommatidium are electrically coupled via gap junctions, allowing light induced currents in one photoreceptor to flow to the other neuro-ommatidial receptors [16]. This coupling is restricted to adjacent neighbors (i.e. R1 is coupled to R2 and R6, R2 to R1 and R3, etc.). The central receptors bypass the neuro-ommatidia, join the efferent lamina axons and continue on to the next neuropile, the medulla.

Glial cells perform a similar role to that of the pigment cells in the ommatidial layer above, and physically separate the neuro-ommatidial axon terminal rings from one another in the lamina, creating distinct units called cartridges [17]. This glial partitioning also results in electrical isolation of the terminal rings [18]. Current does, however, pass through the glial cells, effectively coupling these compartments as well. The neuro-ommatidia, in likc number to the ommatidia, arc repeated thousands of times across the eye [1].

3 Modeling

The model of a neuro-ommatidium can be formulated in two parts, one to account for the phototransduct ion process and the other to account for charge changes in its components. Figure 3 illustrates the nodes and layers required in this model. Phototransduction is discussed elsewhere [12].

In the equations presented in this section, the layers in the model concerned with charge changes will be referred to by number, starting with one for references to the membrane current layer. Where a node can be related to one of the photoreceptors (i.e. R1 through R6), it is assumed to be designated by the same number as that photoreceptor. When only one node is present in a layer, it is designated as node one. The notation used to refer to model quantities will be of two forms. A quantity of the form Am is a constant. A quantity of the form A_{mnp} refers to a concept associated with processing in the m^{th} node of then* layer in the p* neuro-ommatidium.

Equations to describe intercellular charge changes in R1-R6 photoreceptors and the cartridge ECS can be derived from fig. 4 [12]. The resulting equations for the R1-R6 soma and terminal are

$$V_2\dot{X}_{m2p} = -B_1X_{m2p} - B_2(X_{m2p} - X_{m3p}) + I_{in} \qquad (1)$$

$$V_3\dot{X}_{m3p} = B_2(X_{m2p} - X_{m3p}) - B_3(X_{m3p} - X_{14p}) + B_4 \sum_{i\in C}(X_{i3p} - X_{m3p}) \qquad (2)$$

where $C = \{\mathrm{mod}_6(m)+1, \mathrm{mod}_6(m+4)+1\}$.

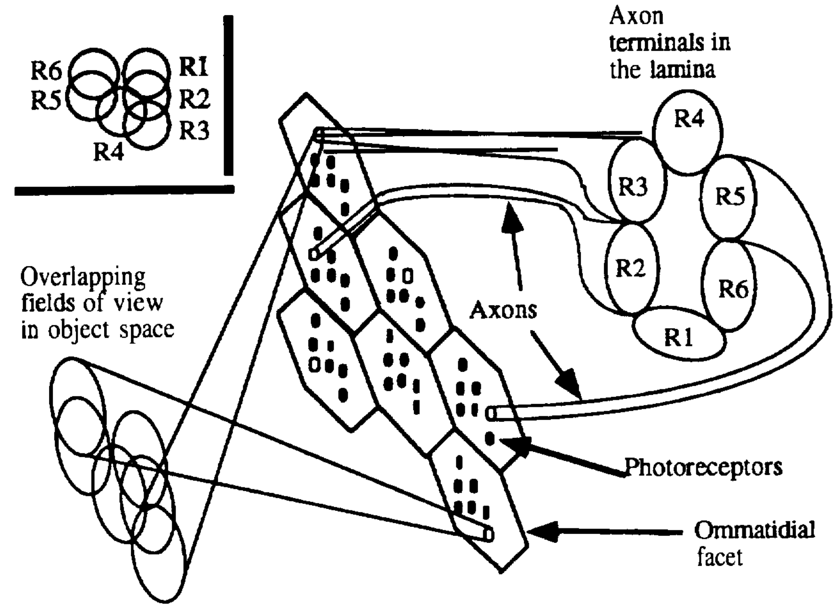


Figure 2 The neuro-ommatidium and the sampling pattern of its photoreceptors. The ommatidial and lamina layers have been flattened. The receptors are depicted as white and black circles within the hexagonal, ommatidial facets. Those receptors in white correspond to the depicted neuro-ommatidium. The R1-R6 axon terminals of these latter receptors join together to form ring-like structures in the lamina. Inset: Every point in space is sampled by six peripheral receptors, each from a different ommatidium. A single R1, R2, etc. is used in each of these overlapped samplings.

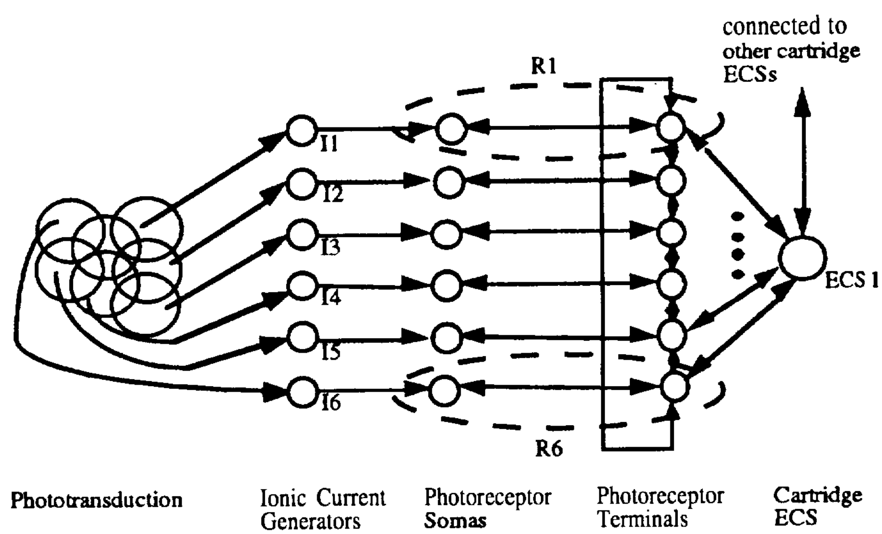


Figure 3 The nodal structure of a model for the flow of charge within the photoreceptors and cartridge ECS of a neuro-ommatidium. The arrows are used to indicate the type of coupling that exists between the nodes.

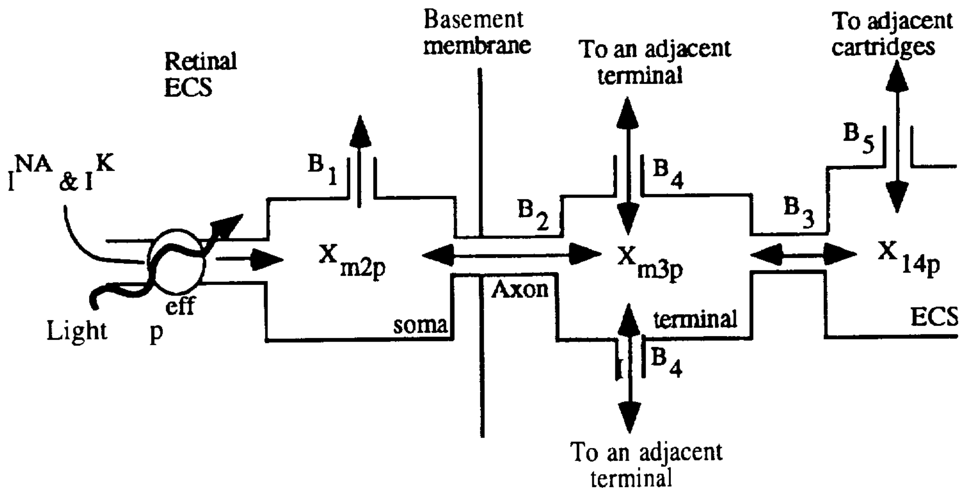


Figure 4 Intercompartmental current flow within an R1-R6 photoreceptor and the cartridge ECS. The-arrows are used to indicate the possible directions for charge flow.

The solutions of eqns. 1 and 2 give normalized, lumped, ionic concentrations (NLICs), relative to a lumped ionic concentration in the retinal ECS, for a receptor soma and a receptor axon terminal, respectively. X_{mip} and Vi represent the NLIC and the volume contained within an element in the i^{th} layer. B_r is a coefficient of flow conductance. I_{in} in eqn. 1 is a lumped input current. The set C in eqn. 2 defines the terminals within the cartridge ring that are adjacent to the m^{th} terminal.

The ionic current I_{in} driving the photoreceptor soma should be predominantly composed of Na^+ and K^+ [13,14,20]. These ions are delivered by three different processes [13,20] that can be combined using the equation

$$I_{in} = -I^{Na}_{m1p} + I^{K}_{bkgnd} + I^{K}_{m1p}. \tag{3}$$

I^{Na}_{m1p} is an inward, light-induced, Na^+ current, I^{K}_{bkgnd} and I^{K}_{m1p} are also inward currents and are defined as the parts of the K^+ current present in an unstimulated photoreceptor and the current above this background level in a stimulated photoreceptor, respectively. Henceforth, references to a K^+ current should be understood to be concerned with only the latter of these two K^+ currents. Na^+ is typically concentrated on the outside of the cell and K^+ on the inside [14]. The increased flow of each into the cell has a different effect, with increased Na^+ flow reducing the lumped, inside-to-outside concentration difference and K^+ flow raising it. With this in mind, the negation of I^{Na}_{m1p} in eqn. 3 does not reflect a difference in current flow direction, but instead its counter role to that K^+ flow.

Taking the expected membrane current properties into account [12], a set of equations of the form

$$\dot{I}^{Na}_{m1p} = -D_1 I^{Na}_{m1p} + \left(D_2 - D_3 I^{Na}_{m1p}\right) * \Gamma \tag{4}$$

$$\dot{I}^{K}_{m1p} = \left(-I^{K}_{m1p} + D_4 I^{Na}_{m1p}\right) * \phi \tag{5}$$

can be used to model the Na^+ and K^+ currents, respectively. The D coefficients of eqn. 4 define the saturation and decay properties of the Na^+ current. D_4 in eqn. 5 defines the steady state K^+ to Na^+ ratio. Γ in eqn. 4 is a transfer characteristic between the effective collected light power in the m^{th} photoreceptor of the p^{th} neuro-ommatidium and

the stimulation that induces the opening of the Na^+ channels. This latter characteristic is dependent on the phototransduction process of the photoreceptor, which records the illumination level via a photochemical reaction [5]. ϕ in eqn. 5 is a function that allows for the adjustment of the K^+decay rate with respect to the lighting conditions. The nature of Γ and ϕ are discussed in [12],

The final equation required for the modeling of the neuro-ommatidial photoreceptors and the cartridge ECS is the one that describe the charge dynamics of the cartridge ECS itself. This can be defined using

$$V_4\dot{X}_{14p} = B_3\sum_{i=1}^{6}(X_{i3p} - X_{14p}) - B_5\sum_{j\in S}(X_{14j} - X_{14p}) . \qquad (6)$$

The set S in eqn. 6 defines the cartridges that are part of the surround of the p^{th} neuro-ommatidium, i.e. the cartridge neighborhood that due to proximity and the glial coupling in the fly's lamina significantly affect the ionic concentration within the ECS of the p^{th} cartridge [4,11,18].

Values for the coefficients required in the model were determined using experimental data and a fitting algorithm [12]. The model is capable of reproducing the responses seen in a fly for wide-field sinusoidal and step stimulation as well as for single photoreceptor illumination [12]. Figure 5 depicts the sinusoidal response properties of the fly and the model. The differences in the phase response are likely due to system delays that were not taken into account in the model's development. Except for this slight discrepancy, the model behavior is quite representative of that seen in the animal.

4. Conclusions

Animal vision offers the potential for the discovery of new approaches to the solution of difficult computer vision problems. The hard-wired nature of insect vision Wows the derivation of a model that captures the functionality of an animal's neurons. The equations presented here model the principles used in the photoreceptor layer of the fly's eye. This model is capable of reproducing the responses recorded in this animal under various types of stimulation.

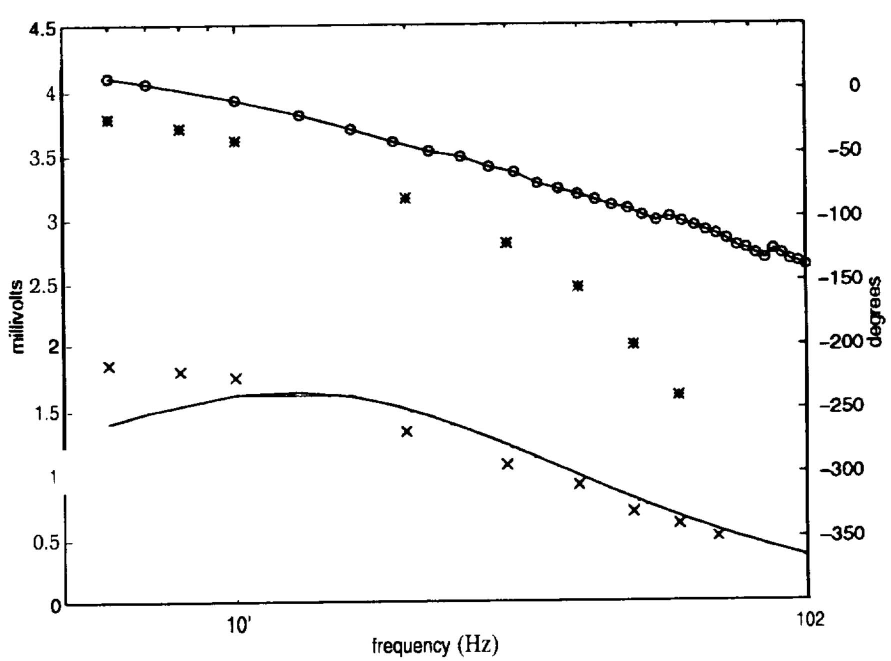


Figure 5 The magnitude and phase of the sinusoidal responses of the model. The details of the stimulation used can be found in [12]. The upper curve (Ø) is the phase response and the lower curve (-) is the magnitude. The data symbolized using "*" and "x" is corresponding experimental results from [7] for fly photoreceptor phase and magnitude responses, respectively. Differences between the model and experimental phase can be attributed to system delays that were not taken into account in the model.

Acknowledgments

This work was performed in the Center for Autonomous Control Engineering at the University of New Mexico. Support was provided in part by NASA under contract NCCW-0087.

References

[1] Braitenberg, V. 1967. Patterns of projections in the visual system of the fly. I. Retina-lamina projections. *Exp. Brain Res. 3, 271-298.*

[2] Collett, T. and Land, M. 1978. How hoverflies compute intercept courses. *J. Comp. Physiol. A* 125, 191-204.

[3] Dill, M., Wolf, R. and Heisenberg, M. 1993. Visual pattern recognition in Drosophila involves retinotopic matching. *Nature* 365, 751-753.

[4] Dubs, A. 1982. The spatial integration of signals in the retina and lamina of the fly compound eye under different conditions of luminance. *J. Comp. Physiol. A 146, 321-343.*

[5] Hardie, R. 1986. The photoreceptor array of the Dipteran retina. *TINS 9,419-423.*

[6] Horridge, G., Mimura, K. and Hardie, R. 1976. Fly photoreceptors III. Angular sensitivity as a function of wavelength and the limits of resolution. *Proc. R. Sot. Lond. B* 194, 151-177.

[7] Järvilehto, M. and Zettler, F. 1971. Localized intracellular potentials from pre- and postsynaptic components in the external plexiform layer of an insect retina. *Z vergl. Physiol. 75,422-440.*

[8] Kirschfeld, K. 1967. Die Projection der optischen Umwelt auf das Raster der Rhabdomere im Komplexauge von Muses. *Exp. Brain Res. 3, 248-270.*

[9] Kuiper, J. 1966. On the image formation in a single ommatidium of the compound eye in Diptera. In *The functional organization of the compound eye, C. G.* Bernhard, ed. Pergamon Press, New York.

[10] Land, M. and Collett, T. 1974. Chasing behavior of houseflies (*Fannia canicularis):* A description and analy-analysis. *J. Comp. Physiol. 89, 331-357.*

[11] Mote, M. 1970. Electrical correlates of neural superposition in the eye of the fly Sarcophagi bullata. *J. Exp. Zool.* 175, 159-168.

[12] Moya, J. 1997. Dissertation. (in preparation)

[13] Muijser, H. 1979. The receptor potential of retinular cells of the blowfly Calliphora: the role of sodium, potassium and calcium ions. *J. Comp. Physiol. A 132, 87-95.*

[14] Nicholls, J., Martin, A. and Wallace, B. 1992. *From Neuron to Brain 3.* Sinauer Assoc., Sunderland, MA.

[15] Pick, B. 1977. Specific misalignments of rhabdomere visual axes in the neural superposition eye of Dipteran flies. *Biol. Cyber.* 26, 215-224.

[16] Ribi, W. 1978. Gap junctions coupling photoreceptor axons in the fist optic ganglion of the fly. Cell *Tiss. Res. 195, 299-308.*

[17] Saint Marie, R. and Carlson, S. 1983. Glial membrane specializations and the compartmentalization of the lamina ganglionaris of the housefly compound eye. *J, Neurocytol.* 12,243-275.

[18] Shaw, S. 1984. Early visual processing in insects. *J. Exp. Biol. 112, 225-251.*

[19] Snyder, A. 1975. Optical properties of invertebrate photoreceptors. In *The compound eye and vision of in-insects, G.* A. Horridge, ed. Clarendon Press, Oxford.

[20] Weckström, M., Hardie, R, and Laughlin, S. 1991. Voltage-activated potassium channels in blowfly photoreceptors and their role in light adaptation. *J. Physiol. 440, 635-657.*

URC97090

12/25

OPTICAL AND ATOMIC FORCE MICROSCOPY CHARACTERIZATION OF PbI_2 QUANTUM DOTS

R. Mu, Y.S. Tung, A. Ueda and D.O. Henderson
Chemical Physics Laboratory, Department of Physics, Fisk University. Nashville, TN 37208

ABSTRACT

Lead iodide (PbI_2) clusters were synthesized from the chemical reaction of NaI (or KI) with $Pb(NO_3)_2$ in H_2O, D_2O, CH_3OH, and C_2H_5OH media. The observation of the absorption features above 350 nm with the help of integrating sphere accessory strongly suggests the quantum dot formation of PbI_2 in solution. Spectral comparison between the synthesized PbI_2 clusters in solution and PbI_2 nanophase by impregnation of PbI_2 in four different pore-sized porous silica indicates that the PbI_2 cluster size in solution is less than 2.5 nm in lateral dimension. Atomic force microscopy (AFM) measurements show that the PbI_2 clusters deposited onto three different molecularly flat surfaces are single-layered. The measured height is 1.0 ± 0.1 nm. The swollen layer thickness can be attributed to the intralayer contraction from the strong lateral interaction among PbI_2 molecules, which is supported by *ab initio* calculation. 'Raman scattering measurement of LO and TO modes of PbI_2 in bulk and in the confined state were also conducted in 50- 150 cm''' region. The observed three bands at 74, %, 106 cm^{-1} are assigned to TO_2, LO_2, and LO, mode, respectively. The relatively small red-shift in LO modes may be caused by the surface phonon polaritons of PbI_2 nanophase in the porous silica.

INTRODUCTION

The study of nanophase materials physically confined in various hosts has attracted much attention in recent years. Depending upon the physical properties of the confining hosts and the nature of the confined materials, many unique mechanical, thermal, and optical properties have been observed and are related to the different types of confinement effects. The physical confinement, interfacial energy at the guest-host boundaries, and the reduction of the confined particle size can lead to, for example, modification of materials' hardness and depression of the melting and freezing transition temperatures. The quantum confinement of the free electrons in metals and the excitons in semiconductors can result in the observation of the colorful surface plasma resonance and the shifts of the band gap. Dielectric confinement due to the confining hosts provides the means to tune the surface plasma resonance frequencies and to enhance the nonlinear optical effects at interface by changing the dielectric constant of the confining hosts. In addition, the observation of the surface phonon polaritons is also due to the dielectric confinement effects. However, it is unfortunate that these fore-mentioned confinement effects are in effect at same time so that a clear fundamental understanding of what exact role of each confinement play is hard to be realized. Further, the often asked question of how these confinement effects modify **the** physical and chemical properties of the confined materials is not at all conclusive. A clear demonstration of the cross-over effects from bulk state to cluster form has not yet been illustrated.

Heavy metal halides, such as PbI_2, form a unique series of layered semiconductor compounds. Besides having potential applications for y-ray detection, the strong intralayer chemical bonding and the weak interlayer van de Waals interaction have made these materials to be good candidates for understanding the cluster formation and growth in confining media and for the study of different types of confinement effects along different crystallographic axes.

Sandroff *et al.*[1] first investigated layered semiconductor clusters in various solvents. Based upon the optical absorption of a solution containing PbI_2 and BiI_3 clusters and the cluster size distribution from TEM measurements, a single layer platelet-like cluster model was proposed. The *disrupt* blue-shifted absorption bands below 350 nm for PbI_2 in different solvents were explained by charge carrier confinement in different-sized crystallite with a single layer thickness. Other research groups[2,3] have re-examined similar systems. Disagreements have been raised in terms of the bands assignments in UV-Vis spectra which may result from the possible I^- or I_3^- presence in solution. However, no efforts have been made to understand how the band gap is truly modified when PbI_2 semiconductor changes its physical dimension from small clusters to its bulk. No solid explanation has been put forward to account for the discontinuous blue-shift of the PbI_2 clusters in solution. In addition, there has been no direct experimental evidence reported in literature to support the proposed disk-like single-layered semiconductor clusters being formed in the solution phase.

Therefore, the motivation of the present research is: i) to study cluster formation and growth mechanisms in the solution phase, which can provide information on crystal growth on earth and in microgravity;[5] ii) to understand how the confining geometry modifies the physical and chemical properties of the restricted clusters in terms of shape, lattice parameters, as well as the thermal and optical properties; iii) to provide direct experimental evidence of the size and morphology of these layered semiconductor clusters formed in solutions; iv) to study quantum confinement effect when PbI_2 is physically confined in four different pore-sized porous glass, which serves as the feed-back to estimate the PbI_2 clusters formed in solution.

EXPERIMENTAL

I. Synthesis of PbI_2 Clusters in Solutions

As reported by others,[1,2] The starting materials are $Pb(NO_3)_2$ and NaI or KI which were purchased from Aldrich with purity of 99'%. No effort was made to further purify the materials. Two reasons why we chose NaI instead of KI in most of our experiments were the following: 1) synthesized PbI_2 clusters are insoluble in pure alcohol but $NaNO_3$ is. Therefore, the dilution of the colloidal suspended solution will not result in much of the alteration of the formed cluster size; 2) in order to unambiguously identify PbI_2 cluster formation and cluster sizes with AFM, the complete removal of $NaNO_3$ from the surfaces and maintaining the minimal perturbation to the PbI_2 clusters are critical so that SPM images can directly reflect the structural and frictional information on PbI_2 clusters themselves. Both $Pb(NO_3)_2$ and NaI (or KI) were prepared in same solvent with known amount of molar concentration. Lead iodide clusters were synthesized by simply mixing the two solutions in the optical cell before subject to various measurements.

II. Impregnation of PbI_2 in Well-Defined porous Glasses

In order to understand quantum confinement effect on PbI_2 particles, an alternative experiment was planned. That is, to impregnated PbI_2 in well-characterized porous glass. The detailed impregnation procedure has been reported elsewhere. [3]Only a brief outline is given in this paper.

Gelsil porous substrates with 2.5, 5, 10, and 20 nm pore diameters were used for PbI_2 impregnation. The PbI_2 was purchased from Aldrich with 99'% purity. The porous glass was first cleaned and then was placed into 10 mm inner diameter and one-end sealed quartz tube. The tube was placed into a vertical furnace. The glass was slowly heated to 450 C in 6 hrs and then was cooled down to 110 C. A sufficient amount of PbI_2 powder was loaded into a smaller glass tube with open ends. Then the tube containing PbI_2 was transported into the bigger tube where the porous glass was dried. The system was heated again very slowly to 450 C. The melted PbI_2 flew out the inner tube and emersed the porous substrates. Upon the 1 hr soaking in the molten state of the PbI_2, the impregnated porous substrate was slowly cooled back to room temperature. The sample tube was properly sealed for characterization.

Hr. Optical Characterization

In order to understand the solution structure and cluster formation kinetics, time resolved and static UV-Vis spectrophotometric measurements were carried out on a Rapid Scan OLIS RSM Spectrophotometer and a Hitichi 3501 spectrophotometer. With OLIS RSM, a series of spectra with time interval of 1-10 ms was collected in the spectral range of 300 - 550 nm. The spectral intensity and frequencies were analyzed as the function of time. The static electronic absorption and extinction spectra were made possible with or without an integrating sphere accessory on a Hitachi 3501 spectrophotometer. All the spectra were measured in 1000-185 mu region with 1 nm resolution.

IV. Atomic Force Microscopy (AFM)

AFM images in TappingMode, constant force, and lateral force modes were obtained with a Nanoscope III atomic force microscope from Digital Instruments. All measurements were performed under ambient conditions with typical relative humidities ranging from 50 - 60%. Well-calibrated E- and A-scanners were used in the present experiments. With A scanner, atomic scale topography images of the hexagonal sheet of SiO_4 in cleavage basal plane of Muscovite mica can be readily obtained on a freshly cleaved mica surface. By comparing the measured value of the nearest neighbor distance of the SiO_4 tetrahedral with the literature value, the lateral resolution of the A-scanner on the atomically flat surface is less than 0.02 nm. Vertical resolution of both A and E scanners in both tapping and constant force modes is less than 0.1 nm. In order to further ensure the consistency in vertical resolution, as we reported eslewhere,[5] a freshly cleaved PbI_2 single crystal surface was imaged which resulted in the observation of half the unit cell distance on c-axis, i.e. a single molecular layer step of PbI_2 along c-axis. The step height was used as an internal height calibration. The measured single layer step height was 0.7 $\pm$0.2 nm which is in good agreement with 0.698 nm in literature.'

V. Raman Measurement

Raman scattering measurements of PbI_2 in its bulk and confined phase in porous silica were conducted with a Spex Raman spectrometer, which is equipped with a double-grating monochrometer and Ti:Saphire laser pumped with Ar^+ laser. The scattered light was collected at 90° to the excitation laser beam. Each spectrum was obtained with 2 cm" resolution and I second integration time. A water cooled photomultiplier tube was used as the detector. The typical excitation energy of the laser was -100 mW at 770 nm.

RESULTS AND DISCUSSION

I. Optical Characterization of So lution Structure

In order to understand solution structure and to identify of possible species present in **solution, efforts were** made to systematically study i) solvent effect **on reactants, i.e.,** NaI and $Pb(NO_3)_2$; ii) solvent effects on the reaction products, i.e., PbI_2 clusters and $NaNO_3$; iii) Pb:I molar ratio effect on reaction products; and iv) concentration effect on the cluster size formation. In the following section, each individual effect will be discussed sequentially.

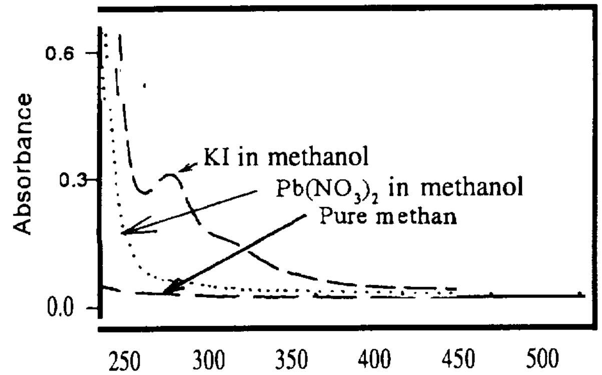


Fig. 1 Optical spectra of NaI **and** $Pb(NO_3)_2$ in methanol solution, and methanol itself.

Fig. 1 illustrates optical absorption spectra of reactants NaI and $Pb(NO_3)_2$ in methnaol solution. Both solvent and $Pb(NO_3)_2$ did not show any absorption band above 250 nm. However, KI solution did show two absorption bands at 270 and 320 urn, respectively. As pointed out by Wang and Herron[3], the observed two peaks at -325 and 270 nm are primarily due to the absorption of I_3^- species in methanol solution. Similar spectra were also observed for NaI or KI dissolved in H_2O, D_2O, propanol methanol solution under ordinary light irradiation. Fig. 2 demonstrates the optical spectra of reaction products by mixing of NaI and $Pb(NO_3)_2$ with different molar concentration ratio. Three different Pb:I ratios were plotted out in this figure. They were Pb:I = 1:2, 1:3, **and 1:4,** respectively. As the number of iodine ions increases, the peaks at 270 mu and 320 urn increase accordingly, while a weak peak at 420 nm remained stationary. This observation further suggests that the bands at 270 nm **and** 320 nm were due to I_3^- complex formed in solution rather than what was believed to be the signature of PbI_2 cluster formation in solution.

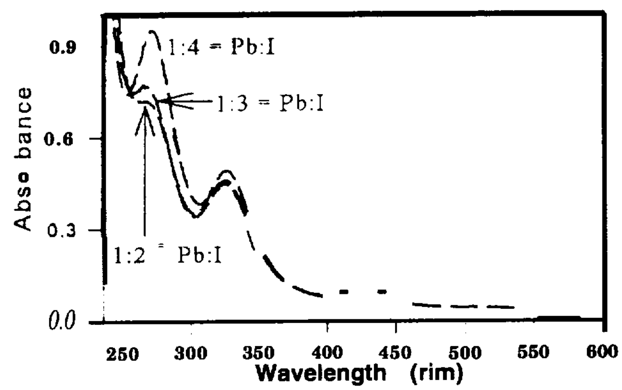


Fig. 2 Optical spectra of the reaction products from mixing of NaI and $Pb(NO_3)_2$ with different molar concentration ratio.

In order to unambiguously identify PbI_2 cluster formation and to clarify true electronic absorption band(s) due to PbI_2 clusters in solution, we have collected both extinction and absorption spectra of PbI_2 colloidal solution at four different concentrations. The extinction spectra were obtained via simple transmission spectra, while the true absorption spectra were collected with the help of integrating sphere accessory. The extinction spectra, consistent with the work reported in the past, [1-3] showed two bands at 270 nm and 320 urn. However, the absorption spectra clearly showed additional bands above 350 mu, which were not observable in extinction spectra, as shown in fig. 3. In addition, the band position shifts to the red and absorption intensity is enhanced as the PbI_2 concentration increases. When the concentration changed from 5.0 to 1.0 mM/l, the electronic absorption band due to PbI_2 clusters blueshifted from 505 nm to 420 mu. At 0.5 mM concentration, the electronic absorption due to the PbI_2 clusters was expected to overlap with the band at 320 Inn.

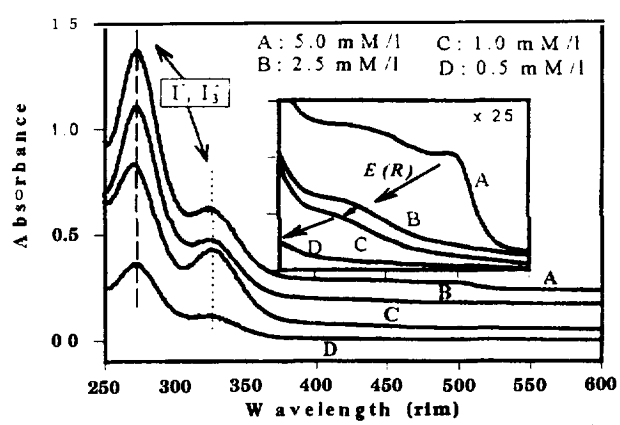


Fig. 3 Absorbance spectra of concentration dependence of PbI_2 in methanol solution,

It is known that when the physical size or dimension of semiconductors is comparable to or smaller than its exciton Bohr radius (for PbI_2, a_B = 1.9 mu), quantum confinement of the electron **and hole will** result in an appreciable blue-shift from its bulk band gap. Based on the effective-mass approximation (EMA) theory proposed by Brus,[9] the absorption energy *E(R)* for a nano-sized semiconductor particle (or quantum dot) can be estimated by the following equation:

$$E(R) = E_g + \frac{\hbar^2\pi^2}{2R^2}\left(\frac{1}{m_e^*}+\frac{1}{m_h^*}\right) - \frac{1.8e^2}{\varepsilon R} + small\ term \qquad (1)$$

where $E(R)$ is the absorption band gap, R *is the size* of the quantum dot, m_e^* and m_h^* are the effective masses of electrons and holes. ε is the dielectric function of the semiconductor. The first term in the equation (1) is the band gap of bulk semiconductor. The second term is due to the quantum confinement of the electron and hole pair. And the third term is Coulomb interaction between the electron and hole. Clearly, the **band gap** of the confined semiconductor is modified by the second ($\propto 1/R^2$) and the third ($\propto 1/R$) terms. Therefore, as the particle size changes, the band gap is changed accordingly. As illustrated in fig. 3, the absorption edge of PbI_2 clusters synthesized in methanol solutions does blueshift gradually as PbI_2 concentration decreases. If it is assumed that the lowest solution concentration has the smallest clusters, then one would expect to observe the largest blueshift in the band gap. Further, as the solution concentration increases, larger clusters are formed that result in an increasing redshift that would ultimately approach the band gap of the bulk material. This trend is consistent with what is predicted by equation (1) and therefore, the absorptions between 550-420 nm for different concentrations are attributed to quantum confined PbI_2 clusters.

I. Confirmation ofthe Existence of Single-layered PbI_2 Clusters Formed in Solution

Fig. 4 shows one of the typical TappingMode AFM images of the PbI_2 colloidal solution deposited on mica surface. It is clear that all PbI_2 clusters have the same height of 1.0 ± 0.1 nm. In order to eliminate any possible structure alteration due to the mica surface, which can lead to false height information, the same PbI_2 colloidal solution was used to deposits on graphite and SAM derivatized mica surfaces. The same height information was also obtained. Therefore, the 1.0 nm height of the image is considered to be the true height of the PbI_2 clusters. It is also conceivable that the PbI_2 clusters must be disk-like as discussed in previous section and proposed by others.[2,7] Then, the PbI_2 clusters on mica surface are orientated with their c-axis perpendicular to the mica surface. Therefore, the measured image height should reflect the layer thickness of the PbI_2 clusters. Due to the tip convolution problem, no attempt has been made to resolve the lateral dimension of the PbI_2 clusters. As pointed out earlier, the measured single layer thickness of the bulk lead iodide is 0.7 nm, while the height of the PbI_2 clusters was 1.0 nm. The 1.4 times layer thickness expansion seems to be surprising. *Ab Initio* calculation of PbI_2 and BiI_3 clusters, on the other hand, has shown that energy-optimized interlayer and intralayer distances of Pb_6I_{12} were 11% expanded and 1 % contracted, respectively, with respect that of the bulk. A geometry optimization calculation suggests that the equilibrium interlayer thickness expanded 40 % while the intralayer distance was contracted by 3 %. Therefore, the observed 1.4 times of the swollen interlayer is expected. As it was argued by Marino *et al.*: In *the case* of Pb_6I_{12} clusters with D_{3h} point group symmetry, the electron-rich I atoms of the top layer are arranged in head-on configuration with respect to I atoms at the bottom layer. For the bulk PbI_2, however, each I atom in the top layer is located directly above the midpoint between two iodine atoms positioned in the bottom layer. Therefore, it was expected that PbI_2 clusters tend to laterally contract and to vertically expand.

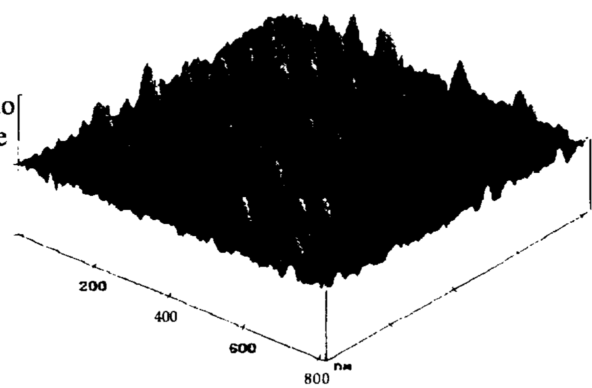


Fig. 4 TappingMode AFM image of lead iodide deposited on a freshly cleaved mica surface. The image height is 1.0 ± 0.1 nm

Optical and AFM characterizations have provided sufficient evidence that the PbI_2 clusters formed in methanol solution is single-layered. The lateral size of the disk-like clusters, based on the quantum confinement of the excitons, is comparable or smaller than that of the exciton Bohr radius $a_B = 1.9$ nm.

III. Optical Characterization of PbI_2 Impregnated in Porous Glass

Fig. 5 showed the electronic absorption spectra of bulk PbI_2 and the PbI_2 physically confined in 2.5,5, 10, and 20 nm pores of silica substrates. As expected, the PbI_2 confined in large pores, the absorption spectrum showed a little blueshift in band gap. However, a noticeable blueshift was observed at ~ 505 nm when the PbI_2 was confined in 2.5 nm pore, which is consistent with the quantum confinement theory proposed by Brus.[10] By comparing the optical spectrum

of PbI_2 impregnated in 2.5 nm in fig. 5 with the spectrum of PbI2 clusters synthesized in methanol solution at 5 mM concentration in fig. 3, a same absorption band edge was observed. This observation provided strong indication that the PbI_2 clusters synthesized in the methanol solution was less than 2.5 nm in size. As the concentration decreases, the cluster size gets smaller so that the absorption band of the PbI_2 clusters was further blueshifted.

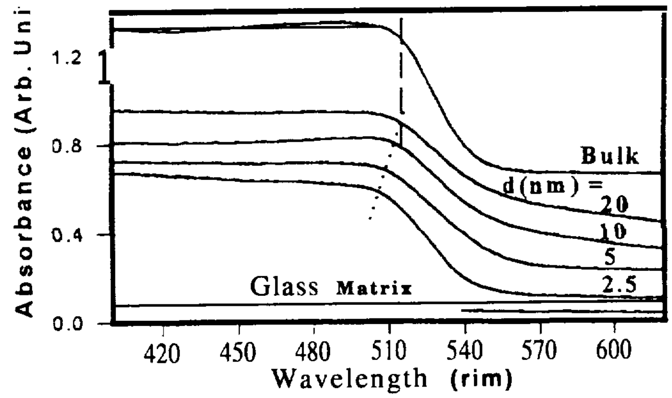


Fig. 5 Electronic spectra of bulk and the confined lead iodide in four different pore-sized silica substrates.

Fig. 6 illustrated Raman spectra of PbI_2 bulk and confined in four different pore-sized silica host. There were three bands observed in 60 - 150 cm-' region. They are at 75, 96, and 106 cm-', respectively, which had been assigned to TO_2, LO_2, and LO_1 optical phonon modes of PbI_2 crystals. As the pore size decreases, LO modes are broadened and the center frequencies seem to be red-shifted. The red-shift of the band at 106 cm^{-2} is more pronounced. It is speculated that the red-shifts in LO modes may be due to the enhancement of the surface phonon modes resulting from the local electric field at guest-host interface for small panicles. However, the experimental results are further from conclusive to make a claim of surface phonon observation in this system. More experiments are current underway to study surface phonon polartons in layered semiconductor quantum dots.

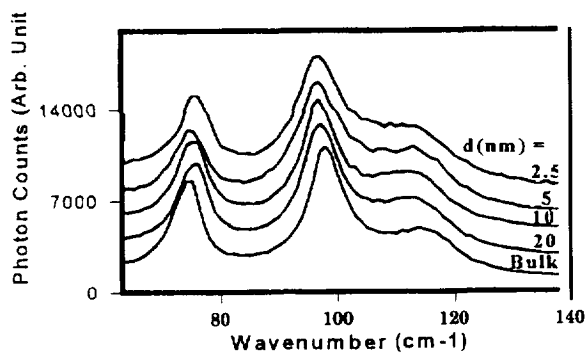


Fig. 6 Raman spectra of the bulk and the confined lead iodide In four different pore-sized silica,

CONCLUSION

Optical characterization of PbI_2 clusters synthesized in methanol solution with integrating sphere has unambiguously confirmed the PbI_2 cluster formation. A spectral comparison of the electronic transitions between the PbI_2 clusters synthesized in solution and physically confined in porous glass suggests that the cluster size in solution is less than 2.5 nm in lateral dimension. This observation is consistent with the concept of strong quantum confinement sets in when the particle size is comparable or smaller than the exciton radius (a_B = 1.9 mu).

AFM measurements of PbI_2 clusters on mica, graphite, and CH_3 surfaces suggest that the clusters formed from solution synthesis are disk-like. A thickness of 1.0 $\pm$ 0.1 nm observed at 0.5 mM concentration confirms, for the first time, that these clusters are single-layered. The 40% expansion of the intertayer distance can be attributed to the finite size effect of the clusters. At small sizes, a strong intralayer chemical bonding can result in the lateral contraction with respect to the bulk value and can lead to the expansion in layer thickness.

LO and TO modes of bulk and the confined PbI_2 were also characterized via Raman scattering measurements. The observed red-shifts and spectral band broadening of LO modes (LO_1 and LO_2) may be attributed to surface phonon polaritons of PbI_2 nanophase in the porous host. However, the experimental results are not yet conclusive.

ACKNOWLEDGEMENT

This work was supported in part by NASA under grant number NAG8-1066 and in part by NASA funded Center for Photonic materials and Devices

REFERENCES

1. M. M. Marine, M. Sawamura, W. C. Ermler, and C. J. Sandroff, Phys. Rev. B 41, 1270 (1990); J. Phys. Chem. 94,7805 (1990); Phys. Rev. B 33, 5953 (1986); J. Chem. Phys. 85, 5337 (1986);
2. A. Roy, D. D. Sarma, and A. K. Sood, Spectrochimica Acts 48A, 1779 (1992);O. I. Mimic, Z. Li, G. Mills, J. C. Sullivan, and D. Meisel, J. Phys. Chem. 91, 6221 (1987).
3. Y. Wang and N. Herron, J. Phys. Chem. 91, 5005 (1987).
4. V. Grasso and G. Mondio "Electronic Structure and Electronic Transitions in Layered Materials", Ed. V. Grasso,

pp -191 (D. Reidel publishing Comp. 1986).
5. For example R. S. Feigelson, J. Crystal Growth 90 1 (1988); F. Rosenberger, J. Crystal Growth 76, 618 (1993).
6. R. Mu, Y.-S. Tung, and D. O. Henderson, Phys. Rev. B [Rapid Comm.] (submitted 1995).
7. Z. K. Tang, Y. Nozue, and T. Goto, J. Phys. Sot. Japan 61,2943 (1992); J. Luminescence 58, 127 (1994); Solid State Commu. **80, 331 (1991).**
8. E. Lifshitz, M. Yassen, L. Bykov, and I. Dag, J. Phys. Chem. **98**, 1459 (1994].
9. D. O. Henderson, R. Mu, A. Ueda, A. Burger, K. T. Chen, and D. O. Frazier, Mat. Res. Sot. Proc. Vol. 366,283 (1995).
10. L. Brus, J. Chem. Phys. 80, 4403 (1984).

5%/43

URC97091

A Remote Sensing Assessment of Land-Use/Land-Cover Changes in The Rio Conchos Watershed of the Rio Grande System, Chihuahua, Mexico

Pedro Muela and Robert Schmidt

Department of Geological Sciences
The University of Texas at El Paso

Introduction

Researchers at the University of Texas at El Paso (UTEP) and the University of Texas at Austin (UTA) are conducting a study of significant land-use/land-cover changes that have taken place within the Rio Conchos watershed area over the past twenty years. Although seldom recognized, the Rio Conchos, which has its headwaters in the remote Sierra Madre Occidental of Chihuahua, Mexico, contributes between 23% and 27% of all the water that flows in the international Rio Grande (Rio Bravo del Norte in Mexico). The quantity and quality of the water flowing from the Rio Conchos is of critical importance to the lower Rio Grande: virtually no water from the upper portion of the Rio Grande flows into the lower half (about 100 miles downstream from the El Paso-Juarez valley the Rio Grande dries up). For all practical purposes, the Rio Conchos is the headwaters for the lower Rio Grande.

The Rio Conchos, which arises at approximately 2,400 meters in one of the wettest and highest areas in the Sierra, flows eastward through a very remote region, and eventually becomes an exotic river as it traverses America's largest arid zone, the Chihuahua Desert, to its junction with the Rio Grande (see Figure 1). The total area of the Rio Conchos watershed is approximately 77,000 square kilometers and is considered to be one of the most threatened river systems in North America. The Conchos river system is under extreme pressure and stress due to improper logging practices, overgrazing. intensification of human use, and other land uses that are frequently environmentally abusive. An assessment of the ongoing and future threat to the environmental integrity of the Rio Conchos promises mutual benefits on both sides of the U.S. - Mexico border and is long overdue.

Project Objectives

The primary objective of this project is to develop a quantitative assessment of land-cover and land-use changes that have taken place over the past two decades within the Rio Conchos watershed and to provide the baseline for monitoring of future changes. We plan to focus our initial research efforts in those parts of the watershed that are currently undergoing the most stress. The main thrust of the research is to conduct an inventory of the land-cover types and land-use activities that can be determined from space, and to evaluate the consequences of observed and expected changes. This project will combine Landsat data with in situ information on climate and hydrology to delineate land uses and identify areas of active erosion within the study area. The area of the watershed requires ten Landsat images for use in this study. Much of the basic information on climate including a 5 megabyte data base already has been compiled by the investigators.

The integration of this information into meaningful and useful statistics, and their spatial characteristics will provide a documentable basis for projecting future impacts and consequences of human activities on the watershed. This study not only represents a much needed baseline, but will serve to educate and to motivate the formulation and/or enforcement of regulations for sustainable multiple use within the watershed. In particular it is essential that the extent and type of land-cover and land-use changes, precipitation distribution including seasonality and variations, reservoir siltation, and watershed efficiency involving seasonality, evaporation, and infiltration be identified and quantified.

Research Methods and Products

1) Inventory and evaluation of previous research. Emphasis will be on the inventory, collection, and evaluation of available data and information from this work.

2) Document climatological and hydrological characteristics of the study area based upon satellite imagery, and ground and aerial reconnaissance.

3) Examine Landsat coverage of the watershed, and determine land-cover, land-use and identify areas of active erosion and possible pollution sources.

4) Measure the storage capacity for Boquillas reservoir. A comparison of the original reservoir dimensional data with that of the present will permit reasonable estimates of erosion rates based upon the added deposition material. This information will be used to relate slop, soil, and land-cover changes to erosion rate.

The product of the research proposed here will be land-usefland-cover maps for two years approximately 20 years apart and land-use and land-cover change maps for that period for portions of the basin with the most sensitive physical setting and undergoing the most rapid change.

Background

Approximately a third of Mexico's major rivers have their origins in the Sierra Madre. More than a dozen of the largest and most valuable irrigation districts in nine states, including Texas, derive a substantial proportion if not all of their water from the Sierra Madre. The Rio Conchos is by far the largest river having its origin in Mexico that reaches the United States. The most important recent threat to the well being of Mexico's most important watershed was a forestry project formulated by the World Bank in the late 1980's. Although the project was promoted as forestry development, it was in effect a logging scheme that was fraught with problems, misconceptions, inconsistencies, and evidence of shortsightedness. Their climatic information, for example, included rainfall and snow accumulations that were far above average and a growing season-tree growth cycle that was very unrealistic. This project was to serve as a model for other logging projects in Mexico. As a result of substantial pressure from environmental and human rights organizations, the project appears to be on hold. Other plans to exploit resources in the watershed will come with or without the World B ank's involvement. Land-use/land-cover information is sore] y needed to provide the framework for analysis of impacts of proposed development plans.

Responses and Consequences

Although most of the study area is considered very remote, the completion of the railroad across the Sierra Madre in 1961, followed by three decades of road building under the GranVision Plan has resulted in very rapid changes in land-cover conversion. The very rough, rugged terrain in the Sierra is largely the result of downcutting into the elevated plateau surface. This has produced five canyons deeper than the Grand Canyon of Arizona. The combination of thin soils, steep slopes and rapid runoff particularly in the upper reaches of the Rio Conchos make the area very fragile to any form of exploitation or development activities. Erosion is already a major problem. Stripping the mountain of the remaining forest would fill the already heavily silted reservoirs and seriously impact, if not eliminate, irrigation agriculture in north-central Mexico and the lower Rio Grande.

Analysis of Remotely Sensed Imagery for Change Detection

Landsat MSS and TM imagery will be the primary sources of remotely sensed data for establishing land- cover/land-use in this study. Although it is only available for recent years, space-based radar imagery will also be evaluated for use in refining future mapping. Additional ancillary and contextual data will also be incorporated where possible in the classification and change detection process. Because the features which are being analyzed differ, analysis of the imagery will require application of an integrated set of approaches for both classification and detection of change. Some areas are comprised of large homogeneous regions of similar spectral response characterized by relatively simple boundary geometry while others will be complex features at multiple scales, consisting of a composite of classes of different sizes. Spectral signatures of multiple targets are often similar at given times of the year, but can be discriminated via multitemporal analysis. Additionally, textural information, ancillary information from field acquisitions, knowledge of class context, and estimates of probable changes in specific classes can dramatically improve results. We propose to utilize a suite of feature extraction and image segmentation and classification algorithms applied at both the region and pixel levels in the study. Specifically, we propose to use a multiresolution feature extracting spatially contiguous features such as areas of deforestation and anthropogenic change in conjunction with a combination of spatial/spectral clustering and supervised maximum likelihood algorithms applied at both a "global" and localized regional level to analyze spectral data. We will also investigate incorporation into our Geographic Information System (GIS).

The proposed region of study is spatially extensive and some classes are large enough to justify use of AVHRR data from the NOAA satellites for multitemporal analysis. We propose to incorporate temporal signatures derived form several years of AVHRR data acquires and temporally reconstructed by the University of Texas Center for Space Research to establish regional level temporal cycles, determine trends in large scale land cover, and resolve problems in vegetation classification. Additional, this sequence will provide data on temporal variation of vegetation signatures and thereby be useful for computing leaf area index (LAI) for rainfall/runoff models.

Finally, because land surface topography plays a central role in hydrology, ecosystem dynamics, biogeochemical cycling, weather, and climate in this area, it will be necessary to utilize topography for land cover classification as well as for determining runoff characteristics and for estimating erosion. Although the U.S. Geological Survey (USGS) has developed digital elevation models (DEMs) for most of the United States (30 m horizontal resolution; 7-15 m vertical resolution), including the state of Texas, a DEM does not exist for the Rio Conchos basin in Mexico.

End-Users of Information From This Project

The remotely sensed imagery, land use and ground cover maps, climatic data, etc., will be placed in a user-friendly Geographic Information System (GIS) in the UTEP Pan American Center for Earth and Environmental Studies (PACES) which has established collaboration with the Center for Health Applications of Aerospace Related Technology. This will allow various universities, organizations, agencies and other end-users to have ready access to the source material and information generated from this project. A GIS format will make the use of this information user--friendly to customers on both sides of the border. It will be especially important to the Border Environmental Cooperation Commission (BECC) as it evaluates projects proposed for funding through NAFTA.

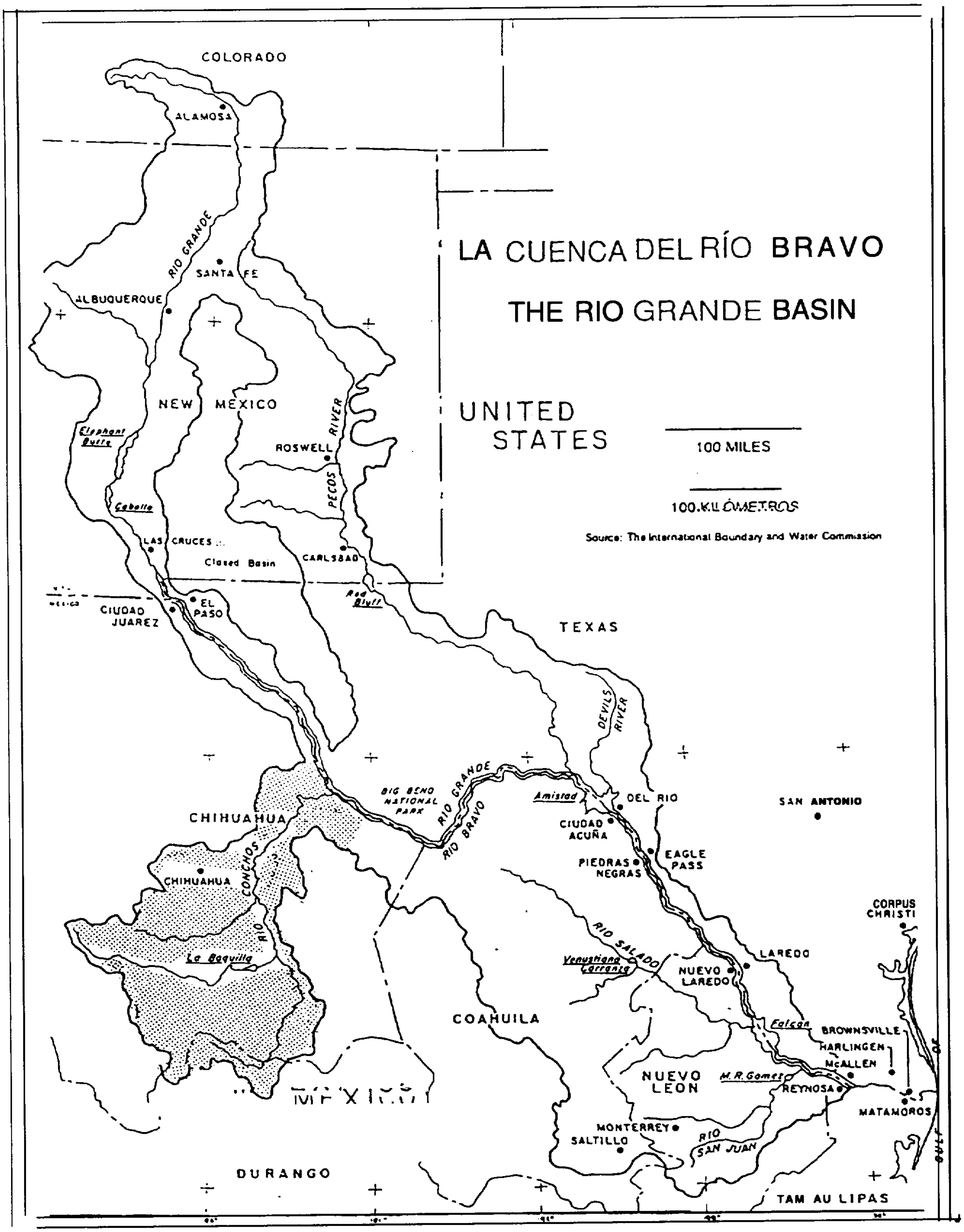


Figure 1. The Rio Grande Basin showing the junction of the Rio Conchos with the Rio Grande (Rio Bravo del Norte in Mexico). The shaded area outlines the Rio Conchos watershed.

N/n 34 592/25

URC97092

AB INITIO STUDY OF THE STRUCTURE AND SPECTROSCOPIC PROPERTIES OF HALOGENATED THIOPEROXY RADICALS

Luis A. Muñoz, R. C. Binning, Jr., Brad R. Weiner", and Yasuyuki Ishikawa*

Department of Chemistry, University of Puerto Rice, PO Box 23346, UPR Station, San Juan, Puerto Rico 00931-3346

1. Introduction

Thioperoxy (XSO or XOS) radicals exist in a variety of chemical environments, and they have as a consequence drawn some interest. HSO, an important species in the chemistry of the troposphere, has been examined both experimentally [1-3] and theoretically [4, 5]. The halogenated (X = F, Cl or Br) peroxy species and isovalent thioperoxy species have been studied less, but they too are potentially interesting because oxidized sulfur species and halogen sources are present in the atmosphere [6]. Learning the fate of XSO and XOS radicals is important to understanding the atmospheric oxidation chemistry of sulfur compounds. Of these, FSO [7, 8] and ClSO [8, 9] are particularly interesting because they have been directly detected spectroscopically.

Recent studies [10, 11] in our laboratory on the photochemistry of thionyl halides (X_2SO; where X = F or Cl) have suggested new ways to generate XSO species. The laser-induced photodissociation of thionyl fluoride, F_2SO, at 193 nm and thionyl chloride, Cl_2SO, at 248 nm is characterized by a radical mechanism [10, 11],

$$X_2SO \rightarrow XSO + X. \tag{1}$$

The structure of FSO has been characterized experimentally by Endo *et cd.* [7] employing microwave spectroscopy. Using the unrestricted Hartree-Fock (UHF) self-consistent field (SCF) method, Sakai and Morokuma computed the electronic structure of the ground $^2A''$ and the first excited $^2A'$ states of FSO [12]. Electron correlation was not taken into account in their study.

In a laser photodissociation experiment, Huber et *al.* identified ClSO mass spectrometrically [13]. ClSO has also been detected in low temperature matrices by EPR [9] and in the gas phase by far IR laser magnetic resonance [8]. Although the structure of FSO is known in detail, the only study, experimental or theoretical, of ClSO has been an *ab initio* HFSCF study by Hinchliffe [14]. Electron correlation corrections were also excluded from this study.

In order to better understand the isomerization and dissociation dynamics of the radical species, we have performed *ab initio* correlated studies of the potential energy surfaces (PES) of ClSO and its isomer ClOS at the QCISD(T)/6-31G* level of theory [15]. For FSO and FOS, more extensive QCISD/6-311G(2df) calculations have been possible, and the results are summarized here.

2. Structure and spectroscopic properties of FSO and FOS in the ground $^2A''$ state

The geometry optimizations and harmonic frequency analyses of the isomers FSO and FOS were done with the Gaussian92 suite of programs [16]. The importance of including d- and f-polarization functions in basis sets for accurate calculation of the equilibrium geometries of fluorine- and sulfur-containing species is documented [17], and adequate sets were employed. The equilibrium geometries and harmonic vibrational frequency analyses of the isomer radicals in the ground $^2A''$ state were determined

* Authors to whom correspondence should be addressed.

by the UHF method with QCISD electron correlation correction based on the UHF reference state. A "frozen core" exclusion of the inner shells from the QCISD calculations was adopted. The optimum geometry of the closed-shell FSO^+ ion was also computed with the same 6-31 1G(2df) basis at the QCISD level in order to obtain an accurate estimate of the adiabatic ionization energy of FS O.

Table 1 shows the optimized geometry of ground state FSO determined at the UHF/6-311G(2df) and QCISD/6-311G(2df) levels of theory. The results of our calculations on FSO are to be found, respectively, in the first and third rows. In their UHFSCF calculations, Sakai and Morokuma [12] computed the electronic structures of the ground $^2A''$ and first excited $^2A'$ states of FSO. They employed 4-31 G basis sets and augmented them with s- and p-type Gaussian bond polarization functions. Their results are reproduced in the second row of Table 1 for comparison. To our knowledge, there has been no *ab initio* correlated study on this system prior to our own. Endo *et al.* [7] carried out a microwave spectroscopic study of ground state FSO and determined its geometry and fundamental vibrational frequencies. The fourth row of Table 1 displays the experimental bond lengths and bond angle.

Table 1. UHF–QCISD/6-311G(2df) optimized geometries of ground state FSO and FSO^+

		F-S (Å)	S-O (Å)	$\angle_{F\text{-}S\text{-}O}$ (°)	Total energy (a. u.)
HF	This work	1.564	1.423	107.4	-571.8555
	Sakai[a]	1.560	1,443	107.9	
QCISD	This work	1.604	.454	108.0	-572.5015
Experiment[b]		1.6023	.4523	108.3	
FSO^+					
HF	This work	1.472	1.361	110.8	-571.5045
QCISD	This work	1.510	1.401	111.5	-572.1517

[a]Ref. [12]   [b]Experiment: Ref. [7]

The bond lengths and angle of FSO^+ at the UHF and QCISD levels are displayed in the last two rows of Table 1. Experimental work by Endo *et al.* [7] as well as *ab initio* SCF studies have shown FSO to be a n-radical with a spin-doublet ground $^2A''$ state. The unpaired electron is in the SO antibonding $\pi^*(4a'')$ orbital. FS 0+, absent the unpaired electron in the antibonding orbital, exhibits shortened S–O and F–S bonds.

Accurate estimation of the adiabatic ionization potential of FSO is important to the interpretation of multiphoton ionization experiments. In such experiments, FSO^+ is produced by ionization via a highly excited valence or Rydberg state of FSO, a state which may well have almost the same nuclear configuration as does FSO^+. At the QCISD level the difference in total energies of FSO (–572.50 15 au.) and FSO^+ (–572.1517 a. u.) gives an accurate estimate (probably within ± 0.1 eV with the basis sets employed) of the adiabatic ionization potential of FS O. The computed estimate, 9.52 eV, falls into a range which indicates that two-photon ionization spectroscopy is feasible with commercially available lasers.

Table 2 presents the geometry of isomeric FOS radical, determined at the UHF/6-311G(2df) and QCISD/6-311G(2df) levels. There are no reported experimental data for the species. With FSO, QCISD

correlation correction brings the calculated bond lengths and angle closer to experimental, by 0.04Å and 0.6°, respectively, than they are at the UHF level. In FOS, the effect of electron correlation on geometry is greater. A change of O. 15Å in the F–O bond length is introduced by including correlation. Taking the FSO results as a guide, one expects the QCISD geometry of FOS to be accurate to within 0.01Å and 1°, respectively, in the bond lengths and the angle. The S–O bond is longer and weaker in FOS than in FSO. The local minimum in thc triatomic potential surface which corresponds to FOS lies 83.7 kcal/mole above the global, FSO, minimum. In two other thioperoxy radicals (XSO—XOS; where X = H and Cl) which have been studied, XSO is also more stable than XOS. In HSO—HOS [5], HSO is the mom stable isomer by only 5.4 kcal/mol, whereas ClSO [14] is 42 kcal/mol more stable than ClOS. The uncertainty in the reported values of the enthalpy of formation of FSO leads to a range of values for the reaction $F(^2P) + SO(X^3\Sigma) \rightarrow FSO(^2A'')$. There is a corresponding range of reaction enthalpies reported, from 75 [18] to 86 kcal/mol [19], depending on the origin of the value for dissociation of $FSO(^2A'')$ to $F(^2P)$ and $SO(X^3\Sigma^-)$ employed. The product $F(^2P) + SO(X^3\Sigma^-)$ state is the lowest-dissociation asymptote correlating with the $^2A''$ ground state of FSO or FOS. Assuming the upper limit value to be correct, the local minimum corresponding to the FOS isomer lies only a few kcal/mol below the dissociation asymptote leading to $F(^2P) + SO(X^3\Sigma^-)$.

Table 2. UHF–QCISD/6-311G(2df) optimized geometries of ground state FOS.

		F-O (A)	S-O (Å)	$\angle_{F\text{-}S\text{-}O}$ (°)	Total energy (a. u.)
HF	This work	1.374	1.589	110.9	-571.7135
QCISD	This work	1.526	1.557	110.9	-572.3681

Table 3 displays the computed harmonic frequencies of FSO and the experimentally observed fundamentals [7]. The S–O and F–S harmonic frequencies computed at the QCISD/6-311G(2df) level arc higher by 20 – 30 cm-' than the corresponding experimental values. The experimentally observed S-O stretching frequency in FSO is larger than the frequency (1 148 cm-') of isolated SO. The S–O bond length in FSO is also shorter by 0.03Å than that in diatomic SO (1.481~). The S–O bond *in* FSO is strengthened by the presence of thc electronegative fluorine, which reduces repulsion among the nonbonding electrons on oxygen and sulfur.

Table 3. QCISD/6-311G(2df) harmonic vibrational frequencies (in cm-l) of FSO and FOS.[a]

		ω_1	ω_2	ω_3
FSO	Calculated	1240	791	413
	Observed[b]	1215	763	396
FOS	Calculated	911	491	362

[a] ω_1, ω_2, and ω_3 correspond, respectively, to the S–O and F–S stretching, and the F–S–O bending frequencies of FSO, and, respectively, to the S-O and F-O stretches and the F-O-S bend of FOS.
[b] Experiment: Ref. 7

The FOS harmonic frequencies appear in the last row of Table 3. There is a substantial difference in the S–O stretch of the two isomers; the frequency in FSO is larger by about 300 cm-' than in FOS. This difference is consistent with the finding that the S–O bond in FSO is shorter and stronger than in FOS. The computed F–S stretching frequency in FSO is also larger by about 300 cm-' than the F–O stretch in FOS. The low F–O frequency in FOS indicates that the bond is substantially weaker than the F–S bond in FSO and consistent with the fact that the potential energy surface is flat near the local FOS minimum region.

3. **Structure** and spectroscopic **properties of ClSO and ClOS in their ground $^2A''$ state**

A b initio correlated calculations were performed on the $^2A''$ Cl–S–O potential surface **in** the region of the ClSO + ClOS isomerization in order to understand the energetic of the process. About two hundred QCISD/6-31 G* and QCISD(T)/6-31 G* calculations were performed with the **Gaussian**92 system to map the surface in the region of interest.

Analysis of the computed potential surface has yielded minimum energy structures and spectroscopic properties of ground state ClSO and ClOS. As with the FSO isomers, the optimum geometry of singly ionized, closed-shell $ClSO^+$ was also computed. At the optimum geometry, a single QCISD(T) calculation was performed to more accurately estimate the total energy. Tables 4 and 5 show the QCISD/6-31G* optimized geometries of ground state ClSO, $ClSO^+$ and ClOS.

ClSO, like FSO, is bent with an unpaired electron in arc* orbital. Similar to the case with FSO, removing the unpaired electron in ClSO forms a $ClSO^+$ ion with shorter S–O and S–Cl bond lengths. **Mulliken** population analysis reveals a rather large formal charge **on** the sulfur.

Table 4. QCISD/6-31G* optimized geometries of ground state ClSO and $ClSO^+$. Reported total energies are QCISD(T).

	Cl-S (Å)	S-O (A)	$\angle_{Cl-S-O}$ (°)	Total energy (a. u.)
ClSO	2.0864	1.4957	109.3	-932.26341
$ClSO^+$	1.9564	1.4500	112.0	-931.92706

Population analysis reveals some of the differences in charge distribution between ClSO and ClOS, FSO and FOS, and between the chloro- and fluoro- pairs. In both ClSO and FSO there is a relatively electropositive atom, S, flanked by a quite negative oxygen and halogen. In FSO the halogen is as electronegative as O, whereas in ClSO it is much less so. Charge flows from the sulfur to the oxygen and halogen; in about equal amounts in FSO but more to O than Cl in ClSO. The two molecules are therefore somewhat ionic, with large dipole moments. At the HF level the ordering of the dipole moments is FSO > ClSO > FOS > ClOS. The S–O bond order in ClSO is about 75°/0 greater than the Cl–S. In FSO the S-O bond order is twice that of the F-S bond.

In ClOS and FOS the bond orders of the two bonds in each are equal, The effect on **bonding**, then, of having the most, rather than the least, electronegative atom in the center of each molecule is to reduce the S–O bond from a double to a single bond. There are differences in the actual details of the charge distributions in the two species. In ClOS a very electronegative atom is flanked by two larger

atoms of roughly equal electronegativity, while in FOS there is a relatively electropositive atom at one end of the molecule, with two highly electronegative atoms bonded to each other at the other end. The charge on the sulfur, as the Mulliken gross atomic population, is positive and the same in both molecules. It is reduced in magnitude from the values for FSO and ClSO. In FOS, F and O both carry the same negative charge, smaller in magnitude than the charges on those atoms in FSO, and reflecting the fact that they are bonded to each other. In ClOS the central oxygen carries quite a large negative charge; Cl is positive to only a slightly lesser extent than S.

Table 5. Optimized geometry of ClOS.

	Cl-O (A)	O-S (A)	$\angle_{Cl-O-S}$	Total energy (a. u.)
QCISD	1.7893	1.6331	116.7	-932.19671*

* Total energy computed at the QCISD(T)/6-31 G* level of theory.

Table 6 shows the harmonic frequencies of the two isomers. The S–O stretch in ClOS is about 300cm^{-1} lower in energy than it is in ClSO, exactly as the corresponding frequencies differ in the FSO—FOS pair (see Table 3). The lower frequency reflects the weakness of the S–O bond in XOS relative to XSO which has been cited above. However, the Cl–O stretch in ClOS is about 100 cm^{-1} higher in energy than is the Cl–S stretch in ClSO. In this pair of frequencies the ClSO—ClOS pair reverses the tendency seen in FSO—FOS. The best explanation of the reversal lies in the contrast in atomic electronegativities of the atoms involved in the Cl–S, Cl–O, F–S and F–O bonds discussed above.

Table 6. Harmonic frequencies (cm-]) for the ClSO radical and its isomer ClOS determined at the QCISD/6-31G* level of theory.*

	ω_1	ω_2	ω_3
ClSO	1098	479	294
ClOS	771	602	309

* ω_1, ω_2, and ω_3 correspond, respectively, to the harmonic S–O and Cl–S stretches and the Cl–S–O bend in ClSO. For ClOS, they correspond to the S–O and Cl–O stretching and Cl–O–S bending frequencies.

Acknowledgment

This work has been supported by NASA through the EPSCoR program (Grant No. NCCW-5 6). Two of the authors (RCB and YI) wish to acknowledge the support of the NIH RCMI Center for Molecular Modeling and Computational Chemistry at the UPR. The support of the Air Force Office of Scientific Research (F49620-93-1-0 11 O) is also gratefully acknowledged.

4. References

[1] Ravichandran, K.; Williams, R.; Fletcher, T. R. Chem. Phys. Letters **1994, 217,375.**

[2] **Kendall,** D. J.; O'Brien, J. J. A.; Sloan, J. J.; MacDonald, R. G. Chem. Phys, Letters 1984, 110, 183.

[3] Kakimoto, M; Saito, S; Hirota, E. J. Mol. Spectrosc. 1980, 80, 334.

[4] Goumri, A.; Laakso, D.; Marshall, P. J. Chem. Phys. 1995, 102, 161.

[5] Xantheas, S. S.; Dunning, T. H., Jr. J. Phys. Chem. 1993, 97, 6616.

[6] Plane, J. M. C. in Biogenic Sulfur in the Environment; Saltzman, E, S., Cooper, W. S. Eds.; ACS Symp. Series #393; American Chemical Society: Washington, D. C., 1989.

[7] Endo, Y.; Saito, S.; Hirota, E. J. Chem. Phys. 1981, 79, 1568.

[8] Radford, H. E.; Wayne, F. D.; Brown, J. M. J. Mol. Spectrosc. 1983, 99, 209.

[9] Nishikida, K.; Williams, F. J. Magn. Reson, 1974, 14, 348.

[10] Wang, H.; Chen, X.; Weiner, B. R. Chem. Phys. Letters 1993, 216, 537.

[11] Wang, H.; Chen, X.; Weiner, B. R. J. Phys. Chem. 1993, 97, 12260.

[12] Sakai, S.; Morokuma, K. Chem. Phys. 52 (1980) 33.

[13] Baum, G.; Effenhauser, C. S.; Felder, P.; Huber, J. R. J. Phys. Chem. 1992, 96, 756.

[14] Hinchliffe, A. J. Mol. Struct. 66 (1980) 235.

[15] Pople, J. A.; Head-Gordon, M.; Raghavachari, K. J. Chem. Phys. 1987, 87, 5968.

[16] Gaussian 92, Revision A. Frisch, M.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. Gaussian, Inc., Pittsburgh PA, 1992.

[17] Langhoff, S. R.; Bauschlicher, C. W., Jr. J. Chem. Phys. 1990, 92, 1879.

[18] Herron, J. T. J. Phys. Chem. Ref. Data 1987, 16, 1.

[19] Takacs, G. J. Chem. Eng. Data 1978, 23, 174.

URC97093

An Investigation of the Migration of Africanized Honey Bees into the Southern United States

/// /;

593/51

Hector Navarro
Pan American Center for Earth and Environmental Studies
University of Texas at El Paso
El Paso, TX 79968
e-mail: hectorv@utep.edu

Introduction

It is estimated that *Apis mellifera scutellata*, a honey bee subspecies from Africa, now extends over a 20 million square kilometer range that includes much of South America and practically all of Central America, and recently has been introduced to the southern United States. African honeybees were introduced into Brazil in 1956 by a Brazilian geneticist, Mr. Warwick Kerr. At the insistence of the Brazilian Ministry of Agriculture, in 1957, 26 colonies were accidentally released in a eucalyptus forest outside São Paulo. The swelling front of the bees was recorded as traveling between 80 and 500 kilometers a year. David Roubik, one of the original killer bee team members estimated that there were one trillion individual Africanized/African honey bees in Latin America. An estimate that is thought to be conservative.

Behavioral Characteristics

Honey bees are not native to the Americas and were introduced during the 1600's by settlers, hence the name European Honey Bees, *Apis mellifera ligustica* (EHB). The freed honey bees and their progeny encountered a hospitable dwelling in Brazil and flourished, in contrast to the EHB that did not fare as well **due** to their introduction from a much milder climatic background. The behavioral characteristics that distinguish the two species are swarming and absconding. Swarming occurs when the queen and many hive members break off to forma new colony, and leave the original colony with a young queen, who repopulates the hive. EHB's swarm perhaps once a year, whereas African bees swarm frequently when flowers are abundant. Absconding is when African honey bees gather honey in the nest and abandon the hive en masse and scout for a more suitable locale, in contrast to EHB's which rarely abscond. Since European and Africanized honey bees are members of the same species *A. mellifera*, they interbreed freely. This has enabled the African bee to do well genetically. The original 26 colonies released represented a bottle neck effect for the introduced bees, where available allelic variations would theoretically lead to extirpation, yet the EHB mating with the African honey bee, hybridized to form an Africanized Honey Bee (AHB). This interbreeding has been detrimental to the EHB population, and has provided the African stock with more genetic variability. Bees that carry a genetic code from the African bee subspecies that are found in the nests of the EHB tend to disrupt hive activity because of the swarming and absconding nature of African bees. Evolutionary pressures for *A. m. scutellata*, include apivarous animals such as man, birds, lizards, badgers etc. as well as nectar availability and climatic factors. Another characteristic commonly known by the public is the excessive level of colony defense by the AHB. All honey bees respond to what they perceive as threats to the nest, however, AHB do so more readily and more vigorously than EHB. The honey bees react to an alarm pheromone called an alarm odor, which is triggered by threats such as vibrations and rapid movement. Once a bee stings it leaves behind its intestines which releases the alarm odor. AHB react in higher numbers than do the EHB and at much longer distances. In Mexico over 1,000 stingings, which include 58 humans deaths, and many domesticated animals have been reported up to 1991. It is important to recognize that AHB colonies are unpredictable and have a greater potential for excessive stinging.

Migration Patterns

The migration of AHB from its introduction outside São Paulo to the United States was completed in less than 35 years. Migration of AHB's farmed out in all directions, and AHB's began to attract attention for it's intense nest defense, and frequent nature of absconding and swarming. Scientists are still questioning how far north the AHB will survive in the United States, as they have the ability to survive different climates. One of the first recordings in the U.S. of these insects was on the southern tip of Texas, in October 15, 1990.

Managed and feral honey bee colonies harvest large quantities of nectar and pollen within kilometers of their nesting sites, These bees are highly developed foraging machines that pollinate many types of plants, including vegetables, fruit trees and flowers. They produce honey and one third of our daily diet from crops that are pollinated through entomophily, by honey bees. Impact on American beekeeping and agriculture from the migrating AHB is unpredictable and is dependent on apicultrists being able to be reactive to the AHB introduction into the southern U.S. Beekeeper that adapt to new conditions will be able to do better and prosper. Some suggested responses involve Quarantines of colonies that must now be regulated, Requeening annually to be assured that apiaries are free of African genetic influence, Training of Beekeeping Hobbyist to ensure safety and African free bee nests, and finally Public health and education awareness programs since there is potential danger from AHB's.

The Killer Bee Research Group at the Pan American Center for Earth and Environmental Studies is developing a geographical information system to assist in the monitoring of the northward migration of the AHB. The geographical information system is a computerized database system containing several layers of data related to factors which influence the spread of the AHB. These data layers are in turn referenced geographically to a base map which provides a natural visualization of the complex interactions among factors influencing bee migration. Remotely sensed data from the Advanced Very High Resolution Radiometer is being processed and entered into the geographical information system. A companion work in this conference [4] addresses the technical aspects surrounding the development of the geographical information system.

Economic Impact

Estimates according to the U.S. Department of Agriculture concerning the overall impact of AHB migration on beekeeping and crop production range from \$26 million to \$58 million annually for beekeeping alone, Annual crop losses are estimated at \$93 million. In some parts of Central America, unmanaged AHB reduced honey production of domestic bees by 60 to 70 percent or more, due to competition of available nectar. This will also impact other bee products which include wax, pollination rental of bees and packaged bees.

Summary and Future Work

The distribution of the AHB in the United States depends on many factors such as ultimate levels of hybridization with existing populations, domestic and feral. Where in the U.S. the northern limit of the AHB will lie is still a matter for extensive speculation. Many studies remain to be done where conclusive trends and scenarios can be understood, regarding the potential range of the AHB in the U.S.

Acknowledgment

I wish to thank NASA for its support of this project through the grant NCCW-0089. Also, I wish to extend gratitude to B yron Wood and Louisa Beck of the Center for Health Applications of Aerospace and Related Technologies at NASA Ames Research Center for their assistance in the formulation of this project.

References

[1] Winston, Mark, Killer Bees, Cambridge, MA: Harvard University Press, 1992.

[2] W. G. Hale and J.P. Margham, Biology, New York: Harper Perennial, 1991.

[3] Laidlaw, Harry H., Queen Rearing, Hamilton, IL: Dadant & Sons, Inc., 1950.

[4] Ward, C. *et.al.*, "The Integration of Geographical Information Systems and Remotely Sensed Data to Track and Predict the Migration Path of the Africanized Honey Bee," Proc. of the 1997 NASA URC-TC '97 National Conference on Education, Aeronautics, Space, Autonomy, Earth and Environment, Albuquerque, NM, Feb. 1997.

Page intentionally left blank

/V II ~~ 5~~

URC97094

Analysis of Texture Using the Fractal Model*

William Navas and Ramón Vásquez Espinosa

Abstract

Properties such as the fractal dimension (FD) can be used for feature extraction and classification of regions within an image. The FD measures the degree of roughness of a surface, so this number is used to characterize a particular region, in order to differentiate it from another. There are two basic approaches discussed in the literature to measure FD: the blanket method, and the box counting method. Both attempt to measure FD by estimating the change in surface area with respect to the change in resolution. We tested both methods but box counting resulted computationally faster and gave better results. Differential Box Counting (DBC) was used to segment a collage containing three textures. The FD is independent of directionality and brightness so five features were used derived from the original image to account for directionality and gray level biases. FD can not be measured on a point, so we use a window that slides across the image giving values of FD to the pixel on the center of the window. Windowing blurs the boundaries of adjacent classes, so an edge-preserving, feature-smoothing algorithm is used to improve classification within segments and to make the boundaries sharper. Segmentation using DBC was 90.8910 accurate.

1. Introduction

The idea behind fractal geometry is that a fractal surface or boundary, when examined in finer detail repeats itself. In other words, a part of it resembles the whole but at a different scale, It can be observed that most of the shapes generated by natural phenomena have this property of self-similarity, lead Mandelbrot[1] and others ([5], [6], [7]) to think that they can be modeled better using fractal geometry rather than Euclidean. Clouds, rivers, coastlines and trees are examples naturally occurring fractals. Our analysis is based on the assumption that most textures exhibit self similarity. Thus, the measurement of fractal dimension (FD) can be used as a discriminator, given that the textures have different degrees of roughness,

2. Fractal Dimension

Traditionally we have regarded points, lines, shapes and objects as having O, 1, 2, 3, . . . dimensions. Hausdorff and Bescovitch not only found that there could be fractional dimensions, but they redefined the whole concept. The Hausdorff/Besicovitch dimension is defined as:

$$D = \frac{\log(N_r)}{\log(\frac{1}{r})} \qquad (1)$$

where N_r is the number of copies of the seed and *r is* the size of the copy relative to the seed also known as the scaling factor. This means a fractal surface can have a dimension between 2 and 3, and a fractal curve can have a dimension between 1 and 2. Fractal dimension of surfaces can be estimated using the a variety of methods, some of which are the Blanket Method and the Differential Box Counting Method.

The methods described by Mandelbrot[1] for measuring the length of a fractal curve can be expanded to 2 dimensions case and applied to images. One approach is to estimate the surface area of the texture at varying resolutions. As the resolution, ε, increases, surface area, A(E), also increases. Plotting, A(E) vs. ε on a log-log scale yields a line whose slope is the gradient g, of the fractal dimension. The gradient is defined as:

$$g = D - D_T \qquad (2)$$

where *D is* the fractal dimension, D_T *is* the topological dimension and g is the gradient. If the object is a curve D_T= 1; for a surface D_T=2.

Another approach to measure FD is to quantify how much area a curve occupies; or, for the two dimensional case, how much space a surface occupies at a given resolution.

* This work was partial] y supported by NASA Grant NCCW -0088.

This is can be done effectively by the box counting method in which the size of the box relate to a given scale, r. And the number of boxes, N_r, that contain points of the surface, is related to the area. The slope of the curve generated by plotting $\log(N_r)$ vs. $\log(1/r)$ yields D.

3. The Blanket Method

This method is a two-dimensional derivation of the Sausage Method explained in[1], and it is discussed by Peleg in (2]. We will use this approach to classify textures 3 textures taken from the Brodatz Album[3] and one synthetic and compare them with other images. By measuring the area, $A(e)$ at decreasing resolution (increasing e) and estimating its derivative on a log-log scale,

$$\varepsilon = \frac{d}{d\varepsilon}\left(\frac{\log(A(\varepsilon))}{\log(\varepsilon)}\right) \tag{3}$$

we obtain a set of features which contains information about the surface's self-similarity, and its roughness. Since all surfaces are not fractals, the plot of eq.3 does not yield a straight line. We will take advantage of this fact and use the whole curve as a signature for identifying a particular texture. Each point generated by equation 3 is a feature of the texture that is fed to a minimum distance classifier for identification.

3.1. Feature Extraction

Blankets above and below the surface are defined by:

$$u_e(i,j) = \max\left(u_{e-1}(i,j) + 1, \max\left(u_{e-1}(m,n)_{(m,n)-(i,j)=1}\right)\right)$$

$$l_e(i,j) = \min\left(u_{e-1}(i,j) + 1, \min\left(u_{e-1}(m,n)_{(m,n)-(i,j)=1}\right)\right) \tag{4}$$

The point is to calculate the volume between the two surfaces. One approach is to integrate the difference between u_e and l_e then divide by the thickness which is 2e. But the blankets are not symmetric and this information is lost at integration. A better approach is to measure the upper and lower volumes independently; from the surface to the blanket, If e increases by one, then the area is the difference between the current volume and the one calculated in the previous iteration. The signatures are the slopes of the best fitting line among every three points of the area. For N area measurements the signatures will have N-2 points. Comparisons were done calculating the distances squared between all textures taking into account upper and lower signatures for each textures.

$$D(i,j) = \sum_{\varepsilon}\left[\left(S_i^+(\varepsilon) - S_j^+(\varepsilon)\right)^2 + \left(S_i^-(\varepsilon) - S_j^-(\varepsilon)\right)^2\right] \tag{5}$$

3.2. Blanket Method Results

Table 1 shows results of the minimum distance classifier. Minimum distances are in bold It is evident that the two images having the same texture content are the ones that show the smallest difference between their features. For these textures, the fractal signature could discriminate to which class a given texture pertains. Table 1 shows differences between the features of training (rows), and the features of the textures to be classified (columns).

	hmpaper 1	pigsk 1	weave 1	Synth 1
hmpaper 2	0.0414	0.8504	0.4571	0.4616
pigsk 2	0.5997	0.0589	0.6037	0.3991
weave 2	0.4573	0.7324	0.0255	0.7775
Synth 2	0.4383	0.4027	0.7421	0.0522

Table 1. Distances between all textures.

4. Differential Box Counting

Differential Box counting is an approximation of the Blanket Method. Referring to Eq. 1, N_r, determined as follows: An image of size M x M is scaled down to a size S x S where $M/2 > s > 1$ and s is an integer, and r is S/M. The (x,y) space is partitioned into an (i,j) of size S x S. On each grid (i,j) there is a column of S boxes. We let the maximum and minimum gray levels of the $(i,j)^{th}$ grid fall on the l^{th} and k^{th} box respectively so that:

$$n_r(i,j) = l - k + 1 \tag{6}$$

$$N_r = \sum_{(i,j)} n_r(i,j) \tag{7}$$

Nr is calculated over 3 values of r. D is obtained substituting on eq. 1 and doing a linear regression.

4.1, Segmentation

Experiments show that the Blanket method is less discriminating than DBC so we used DBC for segment ation as described in Chaudhuri[4]. Since FD does not account for directionality nor absolute gray level, four images were generated from the original. They are called: High Gray, Low Gray, Horizontally and Vertically Smoothed. These preprocessed images are called I_2, I_3, I_4 and I_5, respectively. This images are derived from I_1 which is the original image. The features correspond to the FD of I_i, where i = {1,2,3,4,5}.

4.2. Features

The first feature is the fractal dimension of the original image calculated on overlapping windows of size (2w+1)x(2w+ 1). Since the resulting FD will be between 2 and 3, it is normalized by subtracting 2 so that values for feature 1 will be: f(i,j) = FD-2.

For features 2 and 3 the FD of I_1 and 12 are taken. These are defined by the following rules:

$$I_2(i,j) \begin{cases} I_1(i,j) - L_1, & \forall \quad I_1(i,j) > L_1 \\ 0, & otherwise \end{cases} \tag{8}$$

$$I_3(i,j) = \begin{cases} 255 - L_2, & \forall \quad I_1(i,j) > 255 - L_1 \\ I_1(i,j), & otherwise \end{cases} \tag{9}$$

For features 4 and 5 the FD of I_4 and 15 are taken. These are:

$$I_4(i,j) = \frac{1}{2w+1} \sum_{k=-w}^{w} I_1(i, j+k)$$
$$I_5(i,j) = \frac{1}{2w+1} \sum_{k=-w}^{w} I_1(i+k, j) \tag{10}$$

4.3. Methodology

A sliding window of 17x17 pixels scanned the five features of the image in steps of three pixels. The original image was 254x254x256. W=17 was convenient because it can be subdivided into quadrants of 8x8 neighboring a center pixel (i,j). Nr, can be easily calculated for r={ 1/2, 1/4, 1/8). The resulting measurement of D is assigned to the point (i,j) at the center of the window. The training samples taken from the original image were analyzed by DBC and D was calculated for all them.

Once the FD is calculated a feature smoothing algorithm is applied. This reduces the misclassification that occurs on the boundaries between one texture and another caused by sliding window over two different textures. The filter used works on the spatial domain. It uses a window divided into 4 quadrants. The mean of the quadrant that has the smallest variance is the assigned to the pixel at the center of the window. The technique is known as Edge Preserving Noise Smoothing Quadrant (EPNSQ).

Finally the minimum distance classifier compared distances squared between features of the original image and the mean of the features of the training set samples. Figure 1 shows a block diagram of the whole process.

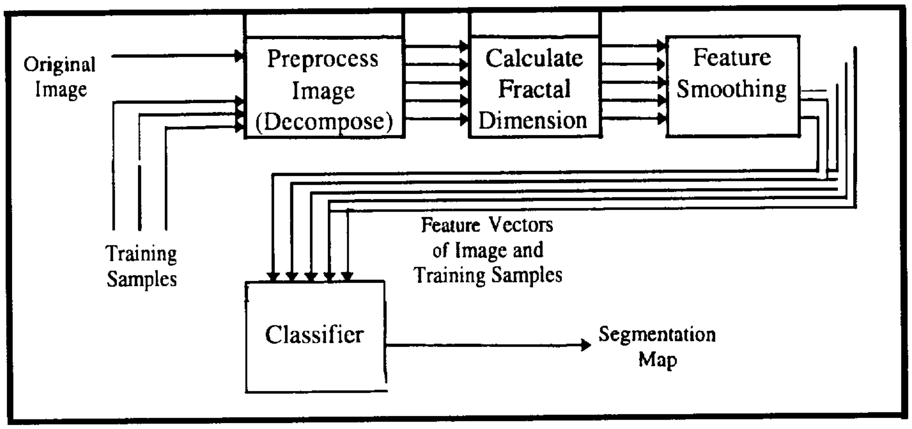


Figure 1. Block diagram of the segmentation process.

4.4. Segmentation Results

The following results were obtained using the method specified on section 4.3. An accuracy of 90.8% on the 3-texture mosaic was achieved. Using only the first feature, the percentage of area correctly classified was 80.3%. To measure accuracy the segmentation map was compared against the actual class map which was crafted by manually specifying the texture class on each sample point.

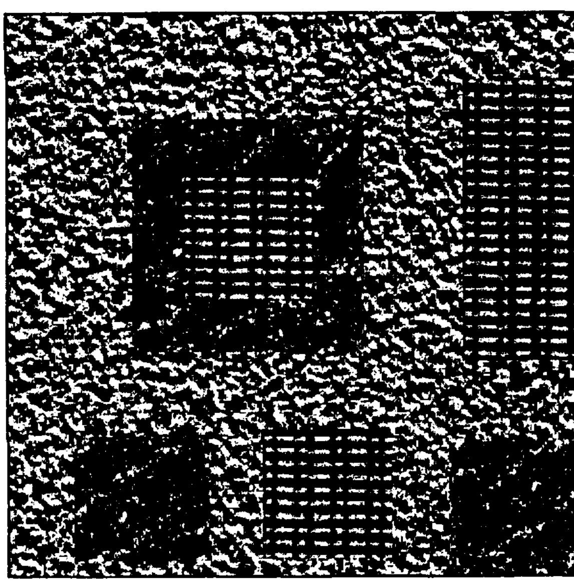


Figure 2. Original Image

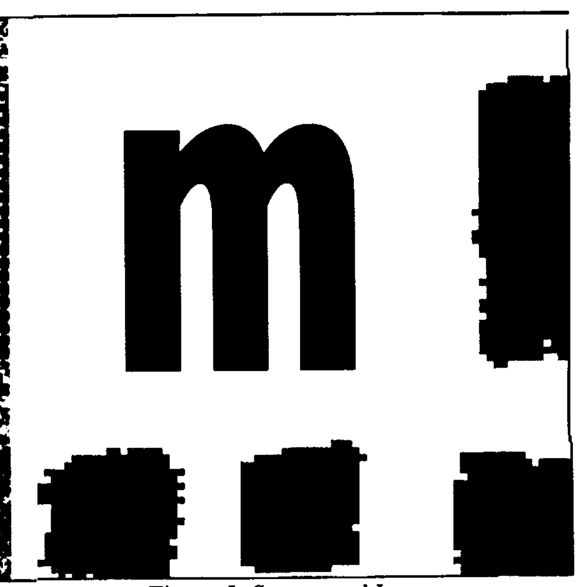


Figure 3. Segmented Image.

The second image used was an aerial photograph, taking samples of the water, vegetation, and agricultural drainage trenches. Psycovisual inspection reveals high correlation between the percieved textures on figure 4 and the segmentation map on figure 5.

5. Conclusion

Through these experiments we could compare the banket method and DBC methods in terms of quality and speed. DB C is faster but segmentation was fair no post-filtering of the features had been applied. The

EPNSQ filter greatly improves classification as it denoises features within a segment while preserving its boundary.

The fractal dimension alone is not enough to characterize texture, On the Blanket Method, upper and lower blankets are used independently as features to account for asymmetry of FD measured on the top side vs. the FD measured on the under side. For DBC, the orignal image was decomposed into low gray, high gray, verticaly smoothed and horizontal y smoothed, so that FD could account for gray-level biases and directionality. This does not assure that two different textures will have the different, but is better than having the original image only. More features can be used at the expense of a linear increase in processing time.

Figure 4. Original Image

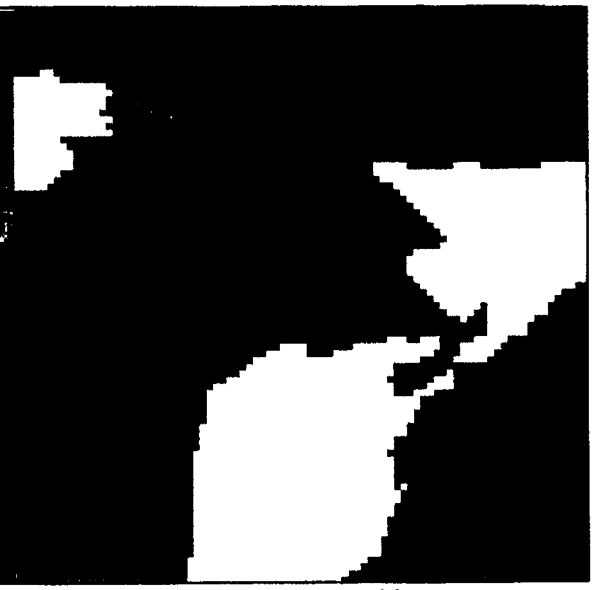


Figure 5. Segmented Image

REFERENCES

[1] B. Mandelbrot - The Fractal Geometry of Nature, W. H. Freeman, NY, 1975

[2] S. Peleg, J. Naor, R. Hartley, D. Avnir - Multiple Resolution Texture Analysis abd Classification, IEEE Transactions, July 1984, PAMI-6 (4), pp. 518-523

[3] P. Brodatz - Textures, A Photographic Album for artists and Designers, NY, Dover, 1966

[4] B. B. Chaudhuri, N. Sarkar, Texture Segmentation Using the Fractal Dimension, IEEE Transactions, PAMI- 17 (l), pp72-77.

[5] R. F. Voss - Random Forgeries, Fundamental Algorithms for Computer Graphics, springer-Verlag, Berlin, 1985) pp.805-835

[6] Lewis F. Jardine, Ian R. Whitworth - Fractal Based Synthesis of Visual Texture, IEEE Colloquium on 'Application of Fractal Techniques in Image Processing, October 1991,

[7] L. M. Linnet, S. J. Clarke, C. Graham, D. N. Langhorne - Remote Sensing of the Seabed Using Fractal Techniques, Electronics and Communications Engineering Journal, October 1991.

Page intentionally left blank

URC97095

What Can Robots Do? Towards Theoretical Analysis

Monica Nogueira

Knowledge Representation Laboratory
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968
email monica@cs.utep.edu

Abstract

Robots have become more and more sophisticated. Every robot has its limits. If we face a task that existing robots cannot solve, then, before we start improving these robots, it is important to check whether it is, in principle, possible to design a robot for this task or not. For that, it is necessary to describe what exactly the robots can, in principle, do.

A similar problem — to describe what exactly computers can do – has been solved as early as 1936, by Turing. In this paper, we describe a framework within which we can, hopefully, formalize and answer the question of what exactly robots can do.

1 Formulation of the problem

The question of "what can be computed" is, basically, solved. One of the fundamental problems of the traditional theory of computing is: *what can, in principle, be computed?* As an answer to this question, Turing developed, in 1936, the notion of a Turing machine: a one-dimensional tape with a bead that moves along it. More complicated computational devices have been proposed since then, devices that use sophisticated memory and sophisticated operations, but whatever can be computed on any of them can still be computed on a Turing machine (slower, but still computed). In this sense, Turing machine does a pretty good formal answer to the question "What can, in principle, be computed?" (for details, see, e.g., [2]).

A new question: what can be done? In the beginning, computers were mainly used for computing. Even when the ultimate goal was to change something in the world, the computer would only generate the instructions, and then these results will be used in the actual control.

Nowadays, a human being is more and more out of this loop. Computers (especially computers employed by robots) are directly linked to motors that move the robot and to the actuators that make the robot change the world.

For such computer-equipped robots, the same question appears: what are the limits of their ability? In plain English, what exactly can the robots do'?

What is known and what needs to be done. Although this question seems to be very fundamental, surprisingly, it has been formulated very recently, in [1].

The main reason why this question has never been analyzed before is that robots are mainly done by practical people, in a manner that mixes precise methods and heuristics. Only recently, when robots have become more and more complicated, and we can no longer rely on our intuitive understanding of what they do, robot designers started to appreciate the value of theoretical methods.

in particular, in [1], the question of *what can, in principle, be achieved by a robot is* formulated for a very specific (and probably, reasonably limited) formal theory of robots. As a result, the main emphasis of this paper is not so much on the analysis of what can be clone by a robot, but on checking whether the theory of robots used in this paper is general enough.

So, the question remains: if we allow most. general robots, what can these robots do?

In this paper, we describe a potential framework for answering this question.

2 Towards the desired formalism

The main objective of a robot is to change the state of the world. Therefore, we are interested in knowing which states can be changed to which. In order to describe that, we must first describe, from the robot's prospective, what are th possible states of the world.

First approximation: 2-D world. For simplicity, let us first consider a robot on a surface (e.g., on Earth). In this case, to describe the actual state oft he world, we must describe what happens at each point of this plane. If we fix a coordinate system, then we are interested in knowing. for every pair of real numbers x and y, what exactly happens at a point with coordinates (x, y).

Discreteness **in space.** In reality, all robot's sensors and actuators have only limited spatial accuracy Ax. As a result, from the viewpoint of this robot, there is no way to distinguish between, a point (x, y) and, e.g., a point $(x + \Delta x/2, y - \Delta x/2)$. In other words, a robot cannot distinguish between the events that happen within a square cell of size Δx x Ax. Hence, to describe the state of the world as it is viewed by a robot and as it can be changed by a robot, all we have to do is describe what exactly is happening within each of these cells.

As the first of such cells, we can take a cell centered in O, i e., a cell $(-\Delta x/2, \Delta x/2) \times (-\Delta x/2, \Delta x/2)$. In this case, all other cells have as centers points oft he type $(i \cdot \text{Ax}, j \cdot \Delta x)$ with integer i and j. Thus, instead of the two *real numbers,* different cells can be characterized by two integers, and the state of the Universe can be characterized by describing, for each cell, what is happening in it.

Each cell can have only finitel y many states. How to describe what is happening in each cell? In reality, there are infinitely many possible things that can happen within each cell. However, in reality, each robot has only finitely many sensors, and each sensor has only finitely many different states. As a result, the number of different sensor reading is finite. Hence, from the viewpoint of how a robot sees the world and what this robot can do, we can distinguish between only *finitely many* different states of each cell.

Let us denote the set of all possible states of each cell by Σ.

How to describe. the state of the world? Resulting description. As a result of the above analysis, we can describe the state of the world by describing. for each pair of integers i and j, in which exactly state $\sigma \in \Sigma$ the corresponding cell is. In mathematical terms, the assignment of a state $\sigma \in \Sigma$ to every pair $(i, j) \in Z \times Z$ is called a *function,* Therefore, a state of the world can be described by a *function* from the set Z x Z of all pairs of integers into the set Σ of all cell states.

We are interested in doable tasks. In our analysis, we are interested in *doable* tasks, i .e., in tasks that a robot can perform during a finite amount of time. During a finite period of time, the robot can only change the state of finitely many cells; therefore, when we describe what a robot can do, we can restrict ourselves only to the contents of finitely many cel Is.

In other words, when we describe possible states of the world, it is sufficient to consider only the situations in which only finitely many cells are non-empty, and all the others are empty. [n mathematical terms, we get the following definition:

Definition 1. *Let a finite set Σ be given that* contains *the symbol* Λ. *Elements of this set will be called states of a cell.*

Let Z denote the set of all integers. By a *state of the world* s, we mean a function $s : Z \times Z \to \Sigma$ such *that for all but finitely many pairs of integers (i, j),* $s(i, j) = \Lambda$. *The* value $s(i, j)$ *is called* a *state of the cell* (i, j). *If* $s(i, j) = \Lambda$, we say *that the cell* (i, j) *is empty.*

The state of the robot. To complete the description of the state of the world, we must also describe where the robot is, and what state the robot is in.

A robot is a finite machine; therefore, it can be in only finitely many different states. Similarly to the description of the Turing machines [2], let us denote the set of all possible states of the robot by K. This set must include the *initial state* so in which the robot starts, and the *halting state* h in which the robot reports that the task is done.

The location of a robot at any given moment of time is described by two coordinates i and j.

Definition 2. Let *a finite set K* be *given that contains two elements* s_0 and h. *Elements of the* state K *will* be called *states of the robot;* SO will called the *initial state,* and h will be called the *final state.*

By a configuration, we mean a tuple (s, k, i, j), *where s is a* state *of the world, k is* a state *of the robot, and i and j* are *integers.*

The robot is the only active agent in the world. For simplicity, we assume that the robot is the only active agent in the world.

This assumption immediately excludes situations like a fire in the plant where, in addition to the robot moving things and trying to extinguish the fire, there is another powerful agent — the fire itself — which drastically changes the states of different cells (by burning their contents).

However, in space applications, this assumption is quite reasonable: a rover moving on the surface of the lifeless planet (e.g., on the Moon or on Mars) is, basically, the only active agent there.

Within this assumption, the only changes in the world are the changes that this robot causes.

How can a robot change the state of a cell? In the Turing machine, the changes were easy to describe: the states of each world cell were simply symbols written on a tape. We could always write or delete a symbol.

For a robot, the situation is not so easy: a state of a cell can mean, e.g., that there is a block in this cell, or two blocks stacked on top of each others. There is no way that a robot can simply change the state of a cell from a block in it to "empty": this would mean that a robot can somehow "eliminate" the block. It is even more unprobable that a robot will be able to "create" a new block out of nothing. All it can do is either *transform* the contents of the cell, or move it.

Let us describe the situation accordingly.

States that are simply marks and states that describe physical contents of each cell. First of all, in our description of the state of the cell, we must distinguish between the *marks* that can be easily rewritten, overwritten, etc., and the states that describe the *physical contents* of the cell.

Definition 3. *Let a subset $M \subseteq \Sigma$, be given that contains* **Λ**. *Elements of the set M will be called marks, and elements of the remaining set $P = \Sigma - M$ will be called physical states.*

How a robot changes things. A robot can easily make a mark, but it cannot convert one block into two. To take these restrictions into consideration, we can describe which cell states can be transformed into which. If we can transform a state σ in to a state u', and if we can also transform a state u' into a state σ'', this means that we will be able, in two steps, to transform the state σ into the state u". Therefore, it is reasonable to require that this "transformability" relation is transitive.

We are mainly interested in non-destructive transformations, so it is also reasonable to assume that if σ can be transformed into σ', then we can always "undo" this transformation and transform σ' back into u.

Thus, the "transformability" relation must be an equivalence relation.

Definition 4. *Let $\equiv$ be an equivalence relation on the set P.*

We will extend this relation to the entire set Σ by assuming that $\sigma \equiv \sigma'$ for every two elements u, u' $\in M$.

If $\sigma \equiv \sigma'$, we will say *that a state σ can be transformed into a state u'.*

A robot can move things. To be able to describe how a robot moves things around, we must divide its states into states in which a robot is not carrying anything, and states in which it "piggybacks" some things from the cell.

Ideally, we would like the robot to be able to carry each state intact, but in reality, it may be necessary first to rearrange the things, i.e., to apply one of the possible transformations. Thus, within the set P of all "physical" states of a cell, we must select a subset C consisting of all "carryable" states.

In this case, in order to describe the state of the robot, we must describe two things: in what exactly the state the robot itself is, and what exactly it is carrying. In other words, the set K of all robot states is equal to the set of all pairs (proper state, carry-on).

Definition 5. *Let C be a subset of the set P with the property that every element $\sigma \in P$ is equivalent to one of the elements from the state C. Elements of the set C will be called carryable states.*

We assume that the state K is a Cartesian product K = K_0 x $C \cup$ {Λ}); if a robot is *in a state $k = (q, \sigma)$, we* will say that *it is in a proper state q and carrying an object u.*

When can one state of the world be transformed into another state.

Definition 6. *Let s* and s' *be two* states of *the* world. *We* say that s *can be (in principle) transformed into s'* if *in* both states, *the total* number of physical *cells* (i.e., cells (i, j) for *which $s(i, j) \in P$) is the same, and these cells* can be placed *in one-to-one* correspondence in *such* a way that *each cell state $s(i, j)$ corresponds* to an *equivalent cell state.*

What can a robot do. When a robot is in a ceratin state, it can, first, stay or move. Second, it can change the state of the cell on which it is currently looking. Third, if the cell is empty and the robot is carrying

something, it can download its carry-on onto the cell (maybe, rearranging it along the way). Vice versa, if the cell is empty, then the robot can load its contents on itself (maybe, rearranging it **so** that it will fit on the robot). As a result, we get teh following definition:

Definition 7. Let M be a set consisting *of five elements L. R, F, B, and S that will be called, correspondingly, "move* left", "move *right", "move forward", "move* back", **and** "slay". *By* a *transition junction, we mean a mapping*

$$\delta : K_0 \times C \times \Sigma \to K_0 \times C \times \Sigma \times M$$

that *satisfies the following property: If* $\delta(q, c, \sigma) = (q', c', \sigma', m)$*, then:*

- *either* $c \equiv c'$ and $\sigma \equiv \sigma'$;
- or $c = \Lambda$, $\sigma' \in M$, and $c' \equiv \sigma$;
- or $c \neq \Lambda$, $' = \Lambda$, and $c \equiv \sigma'$.

Main result. The step-by-step transformation oft he robot is defined accordingly (similar to [2]). Now, we can formulate the main result:

THEOREM.

- Let R *be a robot (described by* Definitions *1–7)*. For every state s, *let* **us** denote by F(s) **the** state into *which* this *robot transforms s.* Then, *F is* a *computable mapping from the* set S of the states **of** *the world* into *itself such that for every state s, this state s can be,* **in principle,** *transformed* **into** *F(s).*
- *Let F : S → S be a computable mapping such that for every state s, this* **state s can be,** *in principle,* **transformed into F'(s).** *Then, there exists a robot* **that** *for every* starting *state s, transforms it into F(s).*

Idea of the proof. The technical description of this proof will take too much space, so we will restrict ourselves to describing the main idea of this proof.

Since in every state, only finitely many cells are used, we reserve two potentially infinite areas: one area will serve as a Turing machine tape on which the robot will simulate the computations necessary to decide what to move where, and another wi II serve as a storage area, where objects will be stored one by one, with a possibility to easily access each of them.

Starting with the borderline elements of the configuration s, we take the contents of each cell one by one and carry them to the storage area. After all these elements arc in storage, we carry them back one by one, filling the elements of the new configuration row by row and column by column so that at each moment of time, we will have easy access. (This is somewhat similar to assembling a ready-made home.).

For space applications, we must consider a 3-D problem. For **space applications, especially for applications to the Space Station, we** need robots that operate in a 3-D world.

All our definitions and results can be easily modified for this case: the main modification is that now we need *three* coordinates (i, j, k) (and not two, as before) to identify a cell.

Acknowledgments. This work was part ly supported by the NASA Pan American Center for Environmental and Earth Studies (PACES). The author is thankful to Chitta Baral, Dan Cooke, Michael Gelfond, Vladik Kreinovich, and Scott Starks for their help and encouragement.

References

[1] H. J. Levesgue, "What is planning in the presence of sensing?", *Proceedings of A AA I'96, Annual Conference of the American Associations on Artifical Intelligence, 1996.*

[2] C. H. Papadimitriou, *Computational Complexity*, Addison Wesley, San Diego, 1994.

5-76/31

URC97096

Minority University System Engineering: A Small Satellite Design Experience Held at the Jet Propulsion Laboratory During the Summer of 1996

Miguel Angel Ordaz
e-mail: ordaz@ece.utep.edu & ordaz@banach.math.utep.edu

Department of Electrical and Computer Engineering
University of Texas at El Paso
El Paso, Texas 79968

Introduction

The University of Texas at El Paso (UTEP) in conjunction with the Jet Propulsion Laboratory (JPL), North Carolina A&T and California State University of Los Angeles participated during the summer of 1996 in a prototype program known as Minority University Systems Engineering (MUSE). The program consisted of a ten week internship at JPL for students and professors of the three universities.

The purpose of MUSE as set forth in the MUSE program review August 5, 1996 was for the participants to gain experience in the following areas:

1.Gain experience in a multi-disciplinary project.
2.Gain experience working in a culturally diverse atmosphere.
3.Provide field experience for students to reinforce book learning.
4. Streamline the design process in two areas: make it more financially feasible; and make it faster.

The MUSE program review also provided the primary goals of the MUSE project per university, JPL and students:

1.The university goals were to develop school curricula to reflect system engineering styles so as to keep academia abreast of the changes in industry.
2.JPL's goal is to devise a faster, cheaper and more effective way to design satellites by teaming with academia and industry.
3. The student's goal was to gain experience working in a multi-disciplinary and multi-cultural environment while gaining knowledge of system engineering.

To execute the purposes and goals of the MUSE program, the participants were involved in conceptually designing a satellite, which was named by these same participants as URANIA.

This paper focuses on showing how some students attending UTEP, did during the internship at JPL in order to meet the purposes and goals of the MUSE program as set forth in the MUSE program review. The students participated in the ten week internship held in its entirety at JPL. JPL provided laboratory space for the MUSE participants at one particular location within the laboratory. The classroom provided by JPL was used by the MUSE group as a project central

where all the project activities were discussed and completed with the frequent visits into the different engineering labs at JPL for guidance and advise on the different stages of the project. The project activities included learning about JPL standards and policy, JPL engineering presentations, leadership discussions provided by JPL personnel and the overall activities that went into the conceptual design of the satellite.

A JPL system engineer was always on duty at the MUSE lab providing the proper guidance for the MUSE group in order to achieve the project goals at the proposed timelines. JPL engineers in no way ever took over the project instead these engineers presented suggestions and viable solutions to the different problems and questions encountered by the students and faculty. The students presented their progress, three times during the internship, to reviewing engineers; the final presentation was to an official JPL review board. Each presentation provided the MUSE group with communication experience as well as with technical questions and suggestions directed at improving the quality and extent of the design work.

Program Objectives

Students participating in the internship were handed the MUSE UNISAT Program guidelines (white paper) for informational and planning purposes only for the MUSE Summer 96 workshop at JPL. The guidelines included a description of opportunity, the program objectives, MUSE UNISAT program constraints, guidelines and requirements, proposal submission information and proposal evaluation, selection and implementation. From this white paper the students began to lay the baseline design of URANIA.

URANIA Mission and Science Objectives

URANIA was to be a low earth orbit resource satellite designed to be launched as a secondary payload in order to decrease cost and use off-the-shelf technology in order to decrease design cycle and cost. At the end of the summer program URANIA had the following characteristics:

1. Shape is to be a Hexagonal cylinder of dimensions (30X40X40) cm.
2. Mass- 30 Kg
3. Stabilization- Gravity gradient boom
4. Processor- A80186
5. Memory- 600 Mbit DRAM
6. Power- 17W BOL, 16.5 W EOL
7. Mission Life- 1 year required and 3 year goal
8. Orbit- Inclination 42°
   Altitude 1100 Km
   Duration 112 min

The science payload instruments for URANIA are as defined in the white paper a cloud imager, a Global Positioning System, and an infrared camera. It was this conceptual URANIA that the MUSE participants had developed at the end of the summer program. Also included in the

design were the organization, schedule and cost plan as well as technical proposal report. These were all prepared and completed by the MUSE participants having only the white paper as a guideline but most of all, having the availability of the expertise provided by the JPL engineers and administrators.

The students divided themselves up in a way in which each student participated in three of the nine subsystems of the satellite while choosing to be a cognizant engineer for one of those subsystems. Each subsystem group was attached to one or two JPL engineers, experts in that subsystem, to provide the necessary advice and tools needed to develop the particular subsystem. While each student was busy developing and designing the particular subsystems assignments it was system engineering which was emphasized rigorously by the JPL system engineers since it played the major role in integrating the subsystems of the satellite and the science instruments. Finally, each of the cognizant engineers was responsible for progress presentations on the assigned subsystem.

The following part of the paper will elucidate a student's experience as a participant in the MUSE program and some technical background on his design involvement with the URANIA satellite providing a conclusion on what the learning experience meant to the student.

MUSE Experience Introduction

This summer my major involvement in the MUSE program came as a system engineer for URANIA where I helped maintain the focus of the design work on the various subsystems of the satellite as well as keeping track of the tradeoffs made of each subsystem and finally developing a satellite power profile graph. I was also involved in the development of the thermal control subsystem of URANIA by being the cognizant engineer which meant that I was responsible for presenting the progress of this subsystem throughout the summer. It also meant that I was the one responsible for making sure that guidelines and deadlines were met as well as to provide information about the subsystem whenever necessary to the system engineers, I also participated in the development of the power subsystem of the satellite by providing the power subsystem cognizant engineer with assistance in trying to define the power consumption and power available to the satellite from the given guidelines and system constraints and by generating a solar incident angle graph. As a mechanisms subsystem engineering assistant I provided help in determining the satellite boom design and its role as an energy provider to the satellite. Finally, I also helped generate a technical proposal report.

System Engineer

The system engineering group met several times during the week to discuss deadlines, project alterations, system implications, tradeoffs and subsystem engineering requirements. The deadlines were measured by using the presentation dates as milestones as to when certain expected work should be completed. Project alterations were brought up to compare to a system restriction or guideline and to make sure that implications of such an alteration only altered the system within the constraints already set. If such an alteration or lack of one had to be made then it had to be documented as a tradeoff. For the subsystem engineering requirements a list of

questions and suggestions was gathered from each of the review board presentations. The list was organized and prepared for each of the subsystems and it was then presented to each of the cognizant engineers for execution and answers.

The system engineering group also met periodically with all the members of the subsystems to update and revise information and requirements of each of the subsystems. The meetings were designed to consolidate subsystem information so that each subsystem would know what the other was up to as far as tradeoffs and alterations were concerned.

Finally, I developed a power profile graph of the system by gathering energy requirements and consumption for each of the subsystems. The power profile becomes one the main tools available to the system engineer because it allows for verification of the power used by any system and thus it acts as a guide in deciding for re-allocation of power as needed. I generated the graph using an excel spreadsheet to make calculations from energy data gathered from each of the subsystems. Another tool I developed as a URANIA system engineer was a table for modes of operation of the satellite. Several modes of operation such as launch, spacecraft stabilization and typical orbit were considered for power consumption, this helped the system engineering group determine if such modes of operation were within power constraints of the system.

Thermal Subsystem

The thermal subsystem design was developed under the guidance of Ray Becker a JPL thermal engineer. Becker provided a presentation on basic thermal physics and design considerations. Becker also suggested that a preliminary thermal analysis be completed by using thermal simulation software and developing a spreadsheet which would house the thermal software output file from which temperature data could be extracted. The thermal constraint as provided by the white paper was to make it a passive system in which no active thermal elements such as radiators and heaters would be used for temperature control of the satellite. Instead the thermal analysis would be used to provide for proper placement of the satellite components and correct usage of thermal paints and heating sensors thus providing passive temperature control.

1. TRASYS Implementation

Using TRASYS thermal software a three dimensional space craft figure is created as input using TRASYS commands.

A space craft external surface layout is then considered where the different materials present in the surface will be considered for solar absorptivity and emissivity.

Worst case orbital angles as well as nadir pointing data and spacecraft altitude is also inputted.

TRASYS is now executed and heat fluxes data is generated.

2. Temperature Data Generated

Using an Excel spreadsheet the TRASYS output data is analyzed by manipulating the heat fluxes data and generating spacecraft temperatures for each of the points considered in the TRASYS thermal simulation.

Finally orbital average temperatures are generated and compared against the worst case orbital angles, where the following was true for URANIA :

At an orbital angle of 70° the orbital average temperature was found to be 277K
At an orbital angle of 0° the orbital average temperature was found to be 289K

3. Conclusion and Future Works

The preliminary thermal analysis provided a safe temperature range for operation of the spacecraft, although not complete, the analysis was used by the URANIA group as an indicator that passive temperature control could be achieved.
Left to do is a conduction analysis where the internal components of the spacecraft and their heat fluxes will be considered so that with the combination of the conduction and preliminary thermal analysis data may generated which will provide knowledge for the proper placement of thermal blankets and paints and thermal sensors or component rearrangement.

Power Subsystem

My role in the power subsystem was to provide the cognizant engineer with the power profile data needed so that the various calculations to determine solar cell array and battery type and amount could be determined. I also provided the cognizant engineer with a solar incident graph used to determine the amount and duration of sunlight available for the solar cell array. The power profile data was the same as the one used to generate the power profile graph which was used in the power subsystem calculations to determine energy efficiency of the batteries and solar cell array. The excel spreadsheet used to analyze the TRASYS output file was also used to generate a solar incident graph. Since a three dimensional spacecraft had been generated for the thermal simulation there was flux data available for each side of the figure thus each flux intensity corresponded to a certain solar incident angle.

Mechanism Subsystem

The satellite will contain a minimum amount of mechanisms, this is reduce probability of mechanism failure resulting in elimination of redundant or backup systems. As of the end of the summer program the mechanism proposed for URANIA were the cloud imager lens cover, an actuated pinpuller which would be used release the third mechanism which is the boom. My involvement with this subsystem came in doing research work for the boom and with making suggestions about how to position the boom and make use of its surface area where solar cells could be positioned to provide added power to the spacecraft.

Conclusion

The MUSE program at JPL provided me the opportunity to gain insight and appreciation for system engineering and its vital role in industry. It also allowed me to perform in the same stage with members of other cultures. It was this cultural diversity which provided the primary element of the learning experience by way of each of the members bringing into the group their own ideas, abilities and scientific mannerisms which left the doors wide open for engaging in an all out effort to execute the project from various cultural angles.

Page intentionally left blank

URC97097

Structural Verification of the Space Shuttle's External Tank Super LightWeight Design - A Lesson in Innovation -

Neil Otte
Marshall Space Flight Center

Introduction

The Super LightWeight Tank (SLWT) team was tasked with a daunting challenge from the outset: boost the payload capability of the Shuttle System by safely removing 7500 lbs. from the existing 65,400 lb. External Tank (ET). Tools they had to work with included a promising new Aluminum Lithium alloy, the concept of a more efficient structural configuration for the Liquid Hydrogen (LH2) tank, and a highly successful, mature Light Weight Tank (LWT) program.

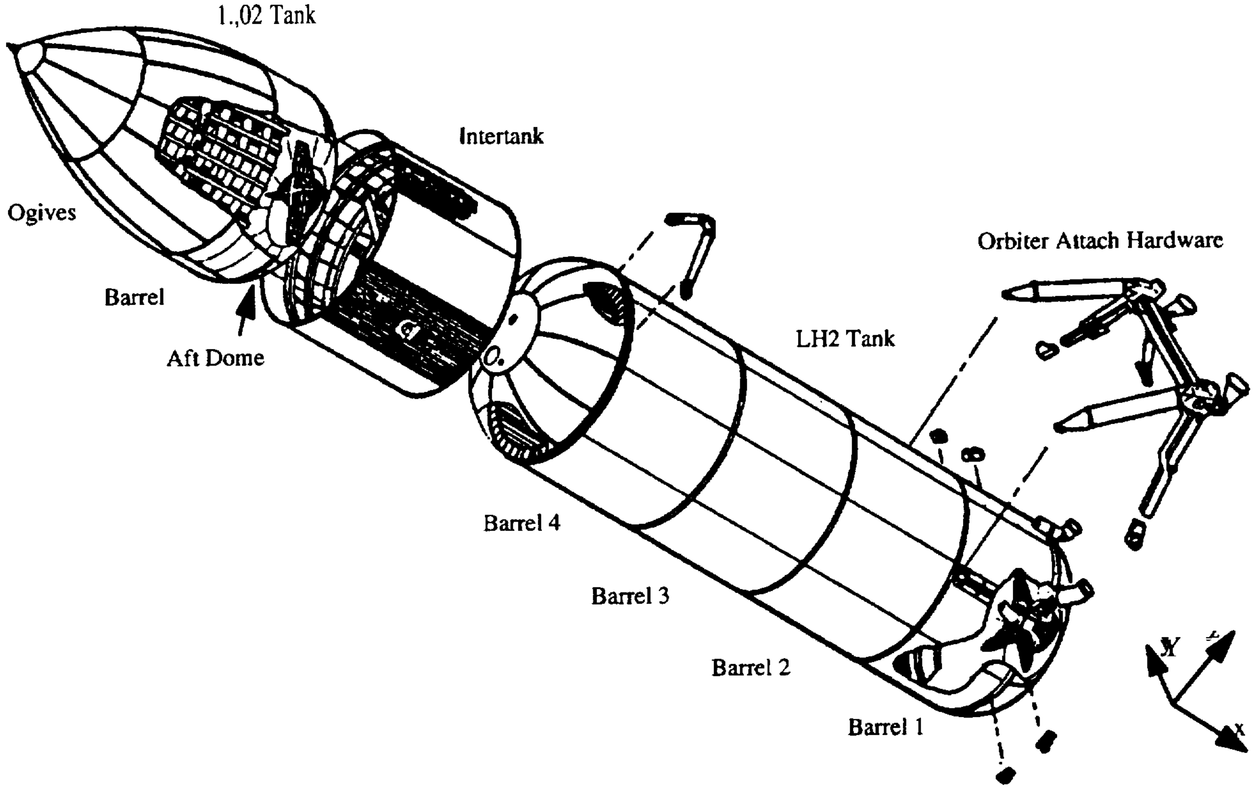


Figure 1. External Tank

The 44 month schedule which the SLWT team was given for the task was ambitious by any measure. During this time the team had to not only design, build, and verify the new tank, but they also had to move a material from the early stages of development to maturity, The aluminum lithium alloy showed great promise, with an approximately 29% increase in yield strength, 15% increase in ultimate strength, 5°/0 increase in modulus and 5°/0 decrease in density when compared to the current 2219 alloy, But processes had to be developed and brought under control, manufacturing techniques perfected, properties characterized, and design allowable generated. Because of the schedule constraint, this material development activity had to occur in parallel with design and manufacturing. Initial design was performed using design allowable believed to be achievable with the Aluminum Lithium alloy system, but based on limited test data. Preliminary structural development tests were performed with material still in the process of iteration. This parallel path approach posed obvious challenges and risks, but also allowed a unique opportunity for interaction between the structures and materials disciplines in the formulation of the material.

While the change from 2219 Aluminum to 2195 Aluminum Lithium for the pressure vessels, i.e. liquid hydrogen (LH2) and liquid oxygen (LO2) tanks, was the primary weight savings tool, it was not sufficient by-itself. Additional changes made in order to achieve the required weight savings included: use of Aluminum Lithium 2090 for the mechanically fastened skin-stringer panels in the Intertank, optimization/machining of the Thermal Protection System (TPS) foam, optimization of the monocoque thickness on the LO2 tank barrels and ogives and the LH2 tank forward and aft domes, optimization of the LO2 aft dome, and, the most significant structural design change, redesign of the LH2 barrel panels from a skin-stringer configuration to an orthogrid stiffened configuration.

Structural Verification Approach

These changes presented the structural verification team with the challenge of defining a structural verification program which protected the flight safety of the Shuttle program, yet met the program's stringent cost and schedule constraints. The team had established the ground rule that all structural verification would be tied to either a test, or flight history of the current LWT. Obviously the technical ideal was a program similar to the original ET verification program, with dedicated structural test articles (STA's) for each major element, i.e. LH2 tank, LO2 tank and Intertank. This approach however, was not feasible given the program's constraints. The question then became, "What tools do we have available for verifying the structure, and how can we make the best use of them?". This question led the team into one of the most innovative structural verification programs ever defined.

The team began by looking at each structural subassembly and their failure modes. An example of this would be barrel 4 of the LH2 tank which exhibits the following critical failure modes: strength in proof, stability of the +Z axis during liftoff and post-staging, stability of the $\pm$ Y axis during liftoff, and stability of the -Z axis during prelaunch. Each of these failure modes had to be verified by a test. As the team worked their way through the structural subassemblies, a program took shape which included a variety of elements: subassembly component testing, maintaining Standard Weight Tank (SWT) or LWT thickness and design, independent structural analysis, proof testing of the LO2 and LH2 tanks, protoflight testing of the LH2 tank, and testing of a dedicated Aluminum Lithium Test Article (ALTA).

Existing Data Base and Design Ground Rules

The obvious advantage that the SLWT verification team had was the wealth of test and flight data from the original SWT and the current LWT. Although many things were changing for the SLWT, many things remained the same. Outer mold line, interface hardware, and load introduction points were unchanged. And although some loads did increase due to other performance enhancements on the Shuttle System - most notably the change from 104°/0 thrust to 106°/0 for the Space Shuttle Main Engines (SSME's) - the load mix was not significantly changed. The team utilized this data base in developing the verification plan. An example of this was the ground rule for maintaining equal or greater stiffnesses in the LH2 tank ringframes. The testing that was performed on the SWT showed that the ringframes were large enough to enforce nodes, and that general stability of the LH2 tank across ringframes was not a critical failure mode. Although some chords were changed from 2219 to 2195 and some web thicknesses reduced, the chord and frame geometry's were maintained, resulting in equivalent or, taking into account the increased modulus of 2195, greater frame stiffnesses than in the LWT. This allowed the team to concentrate on panel stability of the orthogrid design vs. general instability of the LH2 tank. The team also fell back on the existing data base in areas where a test was not feasible. An example of this was the LO2 tank aft ogive regions critical for flight stability. The only technically adequate test for these regions would be a test of a complete LO2 tank; a prohibitive test from both the cost and schedule standpoint. In these regions the membrane thicknesses were maintained at the LWT thicknesses, even though analysis showed that a reduction in thickness was possible. In this way the capability of the SLWT LO2 tank for this failure mode was maintained at better than the LWT capability because of the increased modulus of 2195.

Subassembly Component Testing

The subassembly component tests allowed the team to target a specific design or material change with a test. An example of a subassembly component test was the Intertank skin-stringer/j oint compression tests. Two of these articles representative of different areas of the Intertank were tested; both having .080" thick 2090 Aluminum Lithium skins, but one demonstrating the .063" thick 2090 formed stringers and the other demonstrating the .080" thick 2024 extruded stringers. The test articles were 137.5" long and 33.24" wide and included the skin panel, five hat stringers, two end chords, and two intermediate frames. The articles were mounted in a test fixture and the aft ends deflected radially .625" to simulate the cryogenic shrinkage of the LH2 interface. The aft end was chilled with LN2 to -320° F and an axial compressive load was applied. The articles were loaded to failure, and in both cases, successfully carried the required ultimate load. In all, 13 different subassembly tests were performed, many with two or more articles of different configurations. Additional examples are: Intertank intermediate frame beaded web tests - one with ,025" thick web and one with .032" thick web, LO2 slosh baffle web test, and a cryogenic environments test which subjected a section of the LH2 barrel panel design to hi-axial stresses at -423° F.

The subassembly component testing also highlights an important philosophy which was implemented on the SLWT structural test program. This was: all structural testing which utilized a dedicated test article would be performed to failure in order to find the true capability of the hardware. Test to failure was considered important for a number of reasons. First it protected the program in the event of loads increases, If load increases were proposed, and analytical margin was available which would allow the higher loads to be accepted, the hardware would still be adequately test demonstrated. Second, the data from the instrumentation and final failure load could be used to correlate and refine analysis methods. This had implications not just for the SLWT program but also for future programs. And third, it was recognized that with new designs, and particularly with a new alloy, failure modes and characteristics may be different, It was deemed important to know not just the load at with the structure would fail, but also the manner in which it would fail. This is particularly true given the laminar nature of the short transverse direction of the aluminum lithium alloys. Strict attention was given to looking for any coupling of failure modes.

Independent Analysis

The ground rule was to base all structural verification of the SLWT on test. However this ground rule was deviated from in two areas. The deviations were allowed on a case by case basis, and with additional requirements imposed. The structures which deviated from the ground rule were required to maintain an analytical factor of safety of 2.0 versus the normally required factor of safety of 1.25 to 1.4 for the SLWT. An additional, independent analysis of the hardware was also required. The two areas which were allowed to deviate from the test ground rule were the LO2 tank aft ogive and barrel in areas critical for unpressurized pre-launch stability, and the aft end of the intertank thrust panel which is critical for staging stability. In both cases the original LWT design and thicknesses resulted in high factors of safety, and the structures team was confident that the thicknesses could be reduced safely, The structure was resized by Lockheed Martin (the External Tank prime contractor) maintaining a factor of safety of 2.0. Independent analyses were then performed by Marshall Space Flight Center (MSFC) for the Intertank thrust panel, and by Langley Research Center (LaRC) for the LO2 tank, to verify the analytical factors. Confidence in the analyses was gained by correlation to test data. The aft end of the Intertank thrust panel was tested to failure under axial loads. This data was used to correlate the models and provided complimentary rationale for verification of the redesign. The LO2 analysis was correlated by analyzing the buckling of the LO2 tank forward ogive which occurred during testing of the original SWT Ground Vibration Test Article (GVTA). The analysis accurately predicted the location of the buckle, at the proper load level with a reasonable imperfection.

Proof Tests

Both the LO2 and LH2 tanks have always undergone proof tests on the SWT and LWT programs, and proof testing will continue for the SLWT program. The LO2 tank proof test is a

room temperature hydrostatic test with the addition of a vacuum under the aft dome to increase the delta pressure on the dome. The LH2 tank proof test is a room temperature GN2 test with mechanical loads applied to the Orbiter and SRB attach points at the aft end of the tank. These loads are reacted by a load head at the front of the tank. Because the proof pressures and loads are determined based upon fracture mechanics considerations, and the strength increase is greater then the fracture toughness increase for cryogenic flight temperatures, the room temperature proof tests result in a strength demonstration above limit load. The verification team determined that the minimum test demonstrated factors of 1.12 for every LH2 tank and 1.17 for every LO2 tank, adequately verified the strength of the pressure vessels.

LH2 Tank Stability Verification

The major change for the SLWT occurred in the LH2 tank. Although the ringframe stiffnesses and the longerons which transfer the Orbiter loads into the tank were unchanged, both the configuration of the barrel panels and the material were changed. The material was changed from 2219 to 2195, and the design from a skin-stringer stiffened structure to an orthogrid. The orthogrid design also varied around the circumference of the tank, with different pocket sizes and rib heights resulting in three basic panel designs. The loading of the LH2 tank also varies; including axial load, bending moment, and shear, as well as concentrated loads from the Orbiter and Solid Rocket Boosters (SRB's). While the strength of the LH2 tank was adequately demonstrated in the proof test, stability of the LH2 tank's various panel configurations with the appropriate load profiles also had to be verified. Two test programs were used to verify the stability of the LH2 tank: protoflight testing of each LH2 tank, and a dedicated Aluminum Lithium Test Article (ALTA).

Protoflight Testing

Protoflight testing consists of two test conditions. One protoflight test case was configured to demonstrate stability of the longeron regions in barrels 1 and 2. The testing is performed in the proof test facility, which has the capability of imposing mechanical loads into the tank at the Orbiter and SRB attach points. As an ultimate demonstration of stability is not possible without over stressing the tank and making it unusable for flight, the testing is performed to 115% of flight limit load. In addition to the protoflight test case for stability of barrels 1 and 2, a test case for stability of the aft dome is also performed. The critical load case for the aft dome is driven by the "pinch" loads induced by the rigidly held SRB's as the LH2 tank is filled and the tank attempts to shrink to a smaller diameter. One hundred and fifteen percent of these "pinch" loads are applied during the protoflight testing. The protoflight stability tests will be performed on every LH2 tank.

The protoflight testing highlights another philosophy incorporated by the verification team. Additional risk to the Shuttle during flight was not an option. Confidence in the structural integrity of the vehicle had to be secured in the verification program. The protoflight testing for the SLWT does represent some additional risk to the program when compared to the LWT testing; that being increased risk of losing the LH2 tank and proof facility. But the risk is confined in the ground testing.

The ALTA

The most innovative element of the SLWT structural verification program was the ALTA. The primary purpose of the ALTA was to demonstrate the ability of the orthogrid panel configurations to withstand the SLWT ultimate loads. However the team pushed far beyond that basic goal.

Figure 2. shows the basic configuration of the ALTA. The barrel is representative of barrels 3 and 4 in the LH2 tank and contains the three basic orthogrid configurations. These orthogrid configurations are also representative of the panels on the -Z axis of barrels 1 and 2 (the opposite side of the tank from the longerons). Therefore, between the protoflight test and the ALTA, the entire LH2 tank is demonstrated. Also included on the ALTA was an aft dome from the LO2 tank. The SLWT team had identified the possibility of additional weight savings from the LO2 tank aft dome if a test could be performed. However, a test of the type needed had never been performed before. The LO2 aft dome is stability critical during the end of flight when the Shuttle is accelerating

at 3 g's and the LO2 surface level is falling through the dome, This induces hoop compression in the dome and drives the critical stability margins. To test the dome to the required ultimate condition required a fluid with a specific gravity of 4.2, or 35 lb./gal. This fluid would have to be pumped into the tank against the stabilizing pressure and then flow back out of the tank. A fluid which would meet these requirements had never been made before. After preliminary conversations with people in the oil drilling industry, the challenge was accepted and a lighter redesigned LO2 tank aft dome was included on the ALTA. The ALTA pressure vessel was completed with a 2219 LWT LH2 forward dome. With the ALTA the team had designed a test article which verified major portions of not just one, but two of the SLWT pressure vessels, as well as accepting the challenge to push the boundaries of test technology.

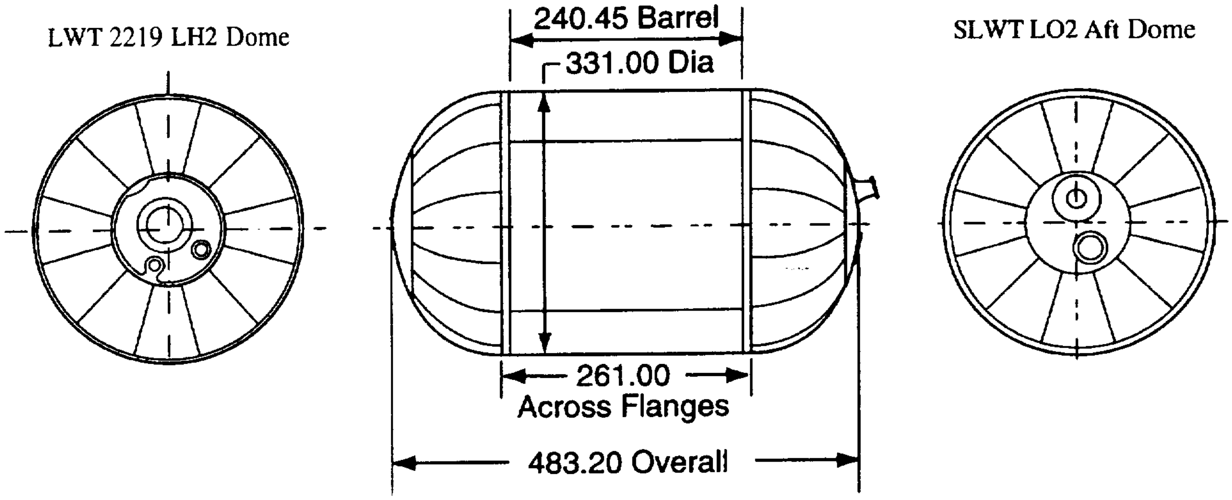


Figure 2. ALTA configuration

The team knew that the program had only one dedicated structural test article, and that they had to glean all the information possible from it. Therefore, the testing of the ALTA was designed not just to qualify the hardware, but to maximize the knowledge gained. The testing started with a pneumatic proof test performed at the Michoud Assembly Facility (MAF) in New Orleans where the ET's are built. The ALTA was then sent to the Marshall Space Flight Center (MSFC) for structural testing. The testing was performed in a refurbished Saturn test stand which was capable of applying axial loads, bending moments, and shears, as well as pneumatic pressurization. The ability to fill the LO2 aft dome with water and the High Density Fluid (HDF) was also provided.

Structural testing was performed over a period of three months and started with the application of influence loads. Axial load, bending moments, and shears were applied one at a time to check that the response of the structure was as predicted, and to verify the instrumentation. A radial load was also applied to the barrel panel at the LO2 feedline support fitting to verify the transverse stiffness of the barrel panel. Application of limit load cases were then started.

As stated previously, barrel 4 of the LH2 tank was subject to 4 failure modes. The strength failure mode had already been tested during the proof test at MAF. The three stability cases were tested on the ALTA at MSFC. By applying the appropriate combinations of axial, shear, and moment, the critical load cases were tested for the +Z, -Z, and $\pm$Y axis panels. First, all three cases were tested to limit flight loads, and then all three were performed to the ultimate flight loads. With the conclusion of the ultimate flight cases the LH2 orthogrid barrel panels - with the exclusion of the barrel 1 and 2 longeron region which are verified in the protoflight testing - were fully verified.

With the verification of the LH2 barrel completed, testing was performed on the LO2 aft dome. The dome was first filled with water to simulate a proof test fill condition. The dome was then tested with the HDF. This test was believed by many to be the riskiest test performed in the SLWT program, not from a SLWT structure viewpoint but from the viewpoint of the ability to perform the test. The Baroid Company was under contract to provide the fluid, pump it into ALTA and then drain it. The fluid used was a mixture of water, a polymer suspending agent, and steel fines. Challenges encountered with the fluid were: the mechanism needed to mix the fluid, the equipment

needed to pump it, and it's ability to flow out of the tank. Working together, NASA, Lockheed Martin, and the Baroid Company met these challenges with the use of concrete trucks, a concrete pump, and the injection of air into the drain line. With these challenges met, the HDF test of the LO2 aft dome was successfully performed. Both the dome and the test set-up performed as expected.

With the completion of the barrel panel tests and the LO2 dome test, the ALTA had performed it's function of verifying the appropriate SLWT hardware. However, the verification team maintained its commitment to gain all the information possible, and to test to failure. In fact, not just one, but three capability tests were performed on the ALTA barrel panels. The liftoff capability test was performed first and the ability of the barrel panel to both carry additional mechanical body loads, and remain stable with decreased pressure was demonstrated, The mechanical body loads were first increased to 125% of the ultimate ,or 175% of the limit, design loads. The body loads were then decreased to the ultimate loads and the pressure decreased. Next, the prelaunch capability test was performed. As the prelaunch case was an unpressurized condition, the case consisted of demonstrating the mechanical loads to 125% of the ultimate, or 162°/0 of the limit, load case. The final capability case was the Post Staging load case. The mechanical loads were first increased to 11 5°/0 of the ultimate, or 146°/0 of the limit, design load condition - the capability of the test stand. The mechanical loads were then decreased to ultimate and the pressure decreased from the design pressure of 31.8 psi. The structure performed linearly until approximately 20 psi, at which time non-linearity was observed in the strain gage readings. Final collapse occurred at ultimate loads and 9 psi pressure. The testing proved the robust design of the orthogrid barrel panels, as well as providing insight into the failure mechanisms of the 2195 Aluminum Lithium. Table 1 shows the load conditions which were tested to capability.

Flight Condition	Design Ultimate Load Condition		Capability Load Condition	
	Pressure (psi)	% of Limit Body Loads	Pressure (psi)	% of Limit Body Loads
Liftoff	17.6	140	17.6 9.6	175 140
Prelaunch	0.0	129.5	0.0	162
Post Staging	31.8	126.5	31.8 20.0 * 9.0 **	146 126.5 126.5

* Denotes approximate condition at which non-linearity was observed in" the gages.

** Denotes condition at which final collapse occurred.

Table 1. ALTA capability conditions

Conclusion

The SLWT verification team was handed the challenge of assuring the adequacy of the SLWT structure to safely perform it's mission. They met that challenge and more. They demonstrated the ability to bring together a verification program utilizing a combination of current data bases, design ground rules, analysis, component testing, protoflight testing, and innovative, cost effective, dedicated structural test articles. They pushed the state of the art of test technology. They looked beyond the current program requirements and expanded the data base in structures technology, They looked to the future and squeezed all the information possible out of each test. They performed their task and more.

URC97098

Force Model for Control of Tendon Driven Hands

Edward Pena
David E. Thompson, Phil

S78/62

NASA Center for Autonomous Control Engineering
Department of Mechanical Engineering
The University of New Mexico
Albuquerque, NM 87131

Abstract

Knowing the tendon forces generated for a given task such as grasping via a model, an artificial hand can be controlled. A two-dimensional force model for the index finger was developed. This system is assumed to be in static equilibrium, therefore, the equations of equilibrium were applied at each joint. Constraint equations describing the tendon branch connectivity were used. Gaussian elimination was used to solve for the unknowns of the Linear system. Results from initial work on estimating tendon forces in post-operative hands during active motion therapy were discussed. Theresults are important for understanding the effects of hand position on tendon tension, elastic effects on tendon tension, and overall functional anatomy of the hand.

Keywords: Artificial Hand, Tendon Driven Manipulators, Hand Biomechanics, Multi-fingered Robotic Hands, Functional Anatomy, Hand Therapy, Clinical Applications, Force Model, Robotics, Prosthetics.

1 Introduction

From studying the functional anatomy of the hand, robotic and prosthetic manipulators have developed to complex designs such as tendon driven manipulators. Although tendon driven hands are complex and difficult to implement, they have many advantages. A tendon driven manipulator allows for a much lighter and streamlined design by placing the actuators in regions more proximal to the hand (e.g. forearm or shoulder area) as opposed to installing them along the joint axes. As a result, more power can be focused on moving the load rather than driving the weight of the motors themselves. The load of the manipulator becomes of greater importance in prosthetics, where the client may be a 50 year old amputee rather than a robotic arm on the assembly line in a manufacturing plant. Those tendon driven systems that make use of elastic elements also offer the advantage of reducing failures such as wear and tear on linkages and joints due to frequent impact of end effecter.

To further designs of artificial hands, a better understanding the of the biomechancis of the human hand must be considered. Therefore, a kinetic model of the human hand is necessary to understand the forces within the hand as well as for control of future artificial hands. If a model existed that predicted the tendon forces that are present when a hand is subjected to external loads, then this hand can now be controlled. By using the model to predict the forces in the tendon for various tasks such as grasping, manipulating, and pushing, the known tendon forces can be used as the input to a neural network and trained to recognize a specific task. After the network is trained, the manipulator can then be controlled for similar tasks as ones used to train the system. For a model to be accurate in predicting the tendon forces in the hand, it is necessary to have the proper biomechanics included. For example, the model must include the proper axes of finger/hand if acceptable grasp is to be obtained. Also, the proper moment arms and considering the elastic property of muscle/tendon unit is very important to insure accurate results of tendon force.

2 Background

The goal of artificial hands in the fields of robotics and prosthetics has always been the same. Each field is still on the quest for the most anthropomorphic hand. Because of stiff design restrictions imposed upon prosthetic hands (e.g. size, weight, functionality, and appearance), the advancement in technology towards a truly anthropomorphic hand has been very slow. In comparison, robotic hands not having those severe design restrictions have risen to a much higher level of dexterity as their hand designs have become more anthropomorphic. These designs contain multi-fingered hands and actuation and power transmission through tendons much like the human hand.

Two tendon driven hands considered the pioneers in the field of robotic hands are the Stanford/JPL Hand (Salisbury Hand) [1] and the Utah/MIT Hand [2]. Although both of these designs showed many anthropomorphic characteristics, they lack anatomically correct joints and tendon configurations; as a result the kinematics and the dexterity of the hand are limited. Recent work on an anatomically correct finger model, the ASU finger [13], emphasizes the importance of human anatomy. This model was used for experiments in motion/control, excursion, and pathomechanics of the finger.

Future work in tendon driven anthropomorphic hands will take advantage of new developments in artificial muscles and advances in kinematics of human hands and arms [7, 8, 9]. A tendon-driven system can naturally make use of advances in artificial muscles such as the ionic polymer gel muscle [10, 11, 12]. With these promising new artificial muscles, the objective of making an artificial hand as anthropomorphic as possible is not far away,

Initial work on a model for estimating tendon forces during active motion therapy has been done [4, 3]. A simplified model was used to predict tendon forces that were generated during post-operative active motion hand therapy. In therapy the patients were required to actively resist a load placed on their finger. The wrist and joints of the fingers axe restricted to movements in the flexion/extension axis of rotation. It is this model that will be presented and discussed in this paper.

3 Development of Model

3.1 Definition of Model Elements

Our model is based on the musculoskeletal anatomy of the human finger, the index finger. This finger contains four bone segments (metacarpal, proximal phalanx, middle phalanx, distal phalanx) and four joints (carpometacarpal (CMC), metacarpophalangeal (MP), proximal interphalangeal (PIP), and the distal interphalangeal (DIP)). There are seven muscle/tendon that control its motion and function, extensor digitorum communus (ED C), extensor indicis (EI), flexor digitorum profundus (FDP), flexor digitorum superficialis (FDS), ulnar interosseous (UI), radial interosseous (RI), and lumbrical (LUM) muscles. Tendons are attached to the end of muscles and run along the surface of bone spanning multiple digits finally inserting into the periosteum (fibrous outer layer of bone). As the tendon is followed distally from the muscle it is found to have a complex arrangement where a tendon may bifurcate into tendon branches or even coalesce into a fibrous sheath as is found in the aponeurosis. Figure 1 shows the skeletal system for the index finger and an example of muscle and tendon complex, the FDP.

3.2 Mechanical Analysis

Different mechanical analogs are used in our model in accordance to the physiologic functions and anatomic constraints of finger joints. The tendon assembly (tendon and tendon sheaths) are modeled as frictionless cable-and-pulley system, and the bone segments in the hand are modeled as rigid links. Each joint is modeled as hinge pins, and since only rotations about the z-axis are considered, modeling only one degree of freedom for the two more proximal joints (CMC and MP) is valid.

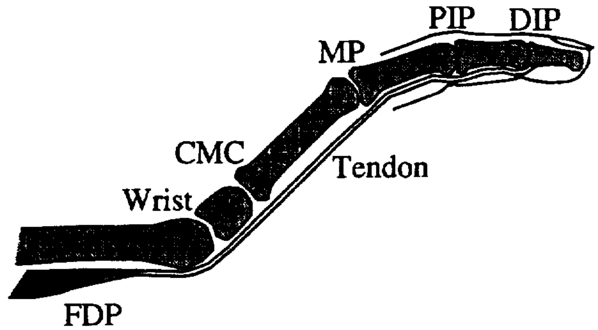


Figure 1: Musculoskeletal Anatomy of Index Finger

There are four joints in our finger model (hindge pins) and therfore, for degrees of freedom. The joints of the finger also called articulations are the functional junction between the digits. The finger is viewed as a kinematic chain and when it is subjected to a load, it is also a closed kinematic chain.

There is a total of five coordinate systems used to define the model, one fixed coordinate system and four moving coordinate systems. The fixed coordinate system is located at an arbitrary position away from the finger. The moving coordinate systems are located at the center of the joint axes for all joints (CMC, NIP, PIP, DIP). A moving coordinate system at each joint is used to easier define the location and orientation oft endon and applied force(s). It also facilitates the ease of defining reaction forces, moments, and relative rot ation between digits. The underlying assumptions and conditions for our model include the following:

1. Normative hand model
2. Force along tendon is constant
3. Moment arms are constant
4. No subluxation of joints during isometric exercise (Stability of joints maintained)
5. Insertion bands of intrinsic muscles/tendon have a single line of action
6. Radial and ulnar lateral bands as well a radial and ulnar interossei are combined using their average moment arm

3.3 Mathematical Model

The entire system is said to be in static equilibrium, implying that each joint must also be in static equilibrium. Therefore, the equations of equilibrium at each joint can be expressed as

$$\sum_{i=1}^{n} \left[\vec{F}_{tendon}\right]_i \cdot \sum_{j=1}^{m} \left[\vec{F}_{external}\right]_j + \vec{F}_{reaction} = 0 \tag{1}$$

$$\sum_{i=1}^{n} [M_{tendon}]_i + \sum_{j=1}^{m} [M_{external}]_j = 0 \tag{2}$$

where $\left[\vec{F}_{tendon}\right]_i$ = tendon force vector

$\left[\vec{F}_{external}\right]_i$ = applied force(s)

$\vec{F}_{reaction}$ = joint reaction force

$[M_{tendon}]_i$ = tendon torque - $\left[\vec{F}_{tendon}\right]_i [r_{tendon}]_i$

$[M_{external}]_j$ = applied torque $\left[\vec{F}_{external}\right]_i [r_{external}]_i$

$[r_{tendon}]_i$ = tendon moment arm

$[r_{external}]_i$ = position vector of $\left[F_{external}\right]_j$ wrt coordinate center

The elastic component of resistance due to soft tissue varies with the joint angle. Llorens [15] approximated this with a three piece elastic model as in Equation 3 and illustrated in Figure 2.

$$k = \begin{cases} 50g - \text{cm/degree} \; \cdots \alpha < \alpha_e \\ 15g - \text{cm/degree} \; \cdots \alpha_e \leq \alpha \leq alpha_f \\ 50g - cm/degree \; \cdots \alpha > \alpha_f \end{cases} \tag{3}$$

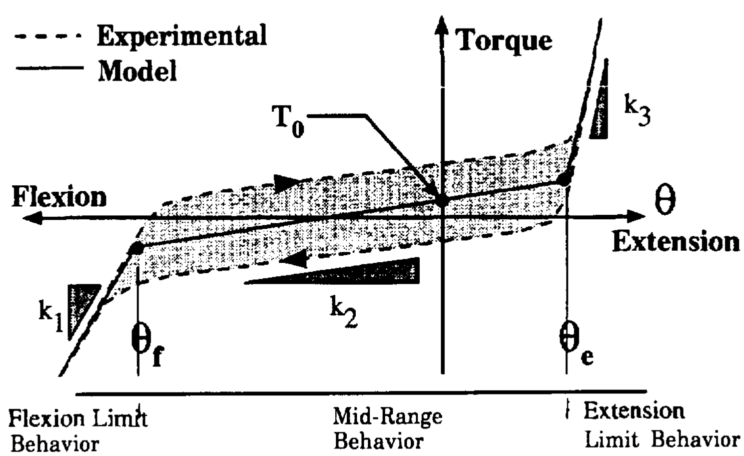


Figure 2: Passive Elastic Torque vs. Angle Curve

The constraint equations are equations that describe how tendon system is interconnected. As an example, in the aponeurosis we find a network of tendons connected together at several points (nodes). Therefore these equations will appear in the form:

Vectoral sum of tendon forces = Vectoral sum of other tendon forces

or as a specific example at the MP-joint:

$$F_{extensor} + F_{interossei} + F_{Lumbrical} = F_{Lateral\ Bands} + F_{Central\ Slip} \tag{4}$$

We end up with n-equations and n-unknowns. The entire system is linear, and can be expressed as

$$[A]x = b \tag{5}$$

where [A] is the coefficient matrix, x is the solution vector of unknown forces, and b is the vector for values of RHS of equation 4. Since the system is linear, we can use the technique of Guassian Elimination, which transforms the coefficient matrix [A] to a lower times the upper triangular matrix, the process called LU decomposition. Once in this form, the unknowns are easily solved using back substitution.

4 Results

There were two cases considered to examine our model: tendon forces with varying joint angle, and tendon forces with the elastic component of soft tissue added. The first case was to demonstrate the effect of finger position on tendon tension. With the wrist extended, the fingers were placed in a nuetral position (no force in tendons) then the fingers were moved into extension only allowing rotation about the MP joint. Calculations were performed with incrementing values of MP joint angle (35-80degrees) to determine the force generated in the EDC. It was found that as the angle increased into extension, the force in the EDC increased significantly from O-g to 1200-g. If the wrist was slightly flexed, the forces in the EDC would decrease because wrist motion influences a change in the length of the flexor tendons resulting in a change of resting angles for the joints.

The second case was considered to demonstrate the elastic effects at the end of motion. An external load of 50g was applied to the middle of the terminal phalanx while the hand was maintained in a tightly flexed fist. This situation allows considerable elastic restoring moments to be generated by large flexion angles. Calculations were performed to determine the flexor tendons (FDS,FDP) forces. The results showed that even for a very small applied load, large forces were generated in the flexor tendons (2050-g FDP and 1650-g FDS). It is important to note that these forces exceed the gapping strength at repair site of the tendon (1600-g).

From the results of these two cases, recommendations for finger loading to insure least potential of rupturing repaired tendons can be made. For example, it is wise not to approach the end limits of range of motion for joints to avoid high levels of generated tendon forces.

5 Conclusions

A two-dimensional force model fort he index finger was developed, and results from initial work on estimating tendon forces in post-operative hands during active motion therapy were discussed. The results showed that there is a significant effect on tendon force for varying joint positions and elastic resistance. It was shown that the magnitude of tendon force in the EDC tendon reached as high as 1200-g under no load for finger extension about MP joint, and the magnitude of force for the FDP tendon reached as high as 1650-g for finger in flexed fist position. With these findings guidelines for active motion therapy were established.

The results are important for understanding the effects of hand position on tendon tension, elastic effects on tendon tension, and overall functional anatomy of the hand. Although it is important to note that this model is very naive because it lacks complexity in certain aspects. One aspect is the assumption of constant moment arms may not be a valid assumption especially for the larger joints (wrist). Another aspect is including a model for the muscle/tendon actuator. Currently, this model defines the tendon force as the total force in the tendon instead of total force being defined as the sum of the tendon and muscle forces. Including the model for the muslce/tendon actuator would account for these components, but it would add another level of complexity to the model. The addition of these aspects to the model will be explored in future work.

An extension for the application of this model is the use of the model to predict required force input of an anthropomorphic exoskeletal structure to reduce fatigue of astronauts during tasks that require hand function. This continues to be an ongoing problem for NASA, and recently there has been an effort made in exploring the use of artificial muscles embedded into the space suit to aid the mobolity of astronauts hands [11, 12].

Acknowledgements

This work was supported in part by NASA under contract # NCCW-0087.

References

[1] Mason, M. and Salisbury, K. 1985. Robot Hands and the Mechanics of Manipulation. Cambridge, Massachusetts: MIT Press, 298pp. Illustrated.

[2] Jacobsen, S., *et al.* 1988. "Design of the Utah/MIT Dextrous Hand." Proceedings of the IEEE International Conference on Robotics and Automation, San Franscisco, California, -10 April 1988.

[3] Evans,R.B., Thompson, D.E. "Application of Force to the Healing Tendon". Journal of Hand Therapy, *Vol.6, No.4*, Oct./Dee. 1992, pp.266-284.

[4] Evans, R. B., Thompson, D.E. "An Analysis of Factors that Support Early Active Short Arc Motion of the Repaired Central Slip." Journal of Hand Therapy, 1993, pp.187-201.

[5] Brand, P.W., Clinical Mechanics of the Hand. St. Louis: C.V. Mosby, 1985.

[6] Chao, E.Y, Opgrande, J. D., Axmear, F.E. "Three-Dimensional Force Analysis of Finger Joints in Selected Isometric Hand Functions". J. Biomechanics, *Vol.19, No.* 6, 1976, pp.387-396.

[7] Agee J, Hollister A and King F. 'The Longitudinal Axes of Rotation of the Metacarpalphalangeal Joint of the Finger,' *J. Hand Surg 11*(A), 1986, p.767.

[8] Buford, WL, Hollister, AM, and Myers, L.M. 'Three Dimensional Computer Simulation of Thumb Carpometacarpal Joint Kinematics,' *Abstracts of the First World Congress of Biomechanics*, Vol.II, pp.161, 1990.

[9] Hollister A, Buford, WL, Myers, LM, Giurintano, DJ and Novick, A. 'The Axes of Rotation of the Thumb Carpometacarpal Joint,' J. Orthopaedic Research, *Vol. 10, No.* 3, pp.454-460, 1992.

[10] Mojarrad, M., Shahinpoor, M., "Ion Exchange Membrane-Platinum Composite Actuator as Electrically Controllable Artificial Muscles", Third International Conference on Intelligent Materials, Lyon, France, June 96.

[11] Meghdari, A., Jafarian, M., Mojarrad, M., Shahinpoor, M., "Exploring Artificial Muscles as Actuators for Artificial Hands", ASME Conference on Smart Materials and Intelligent Systems, Albuquerque, NM Spring 94.

[12] Shahinpoor, M., Wang, G., Mojarrad, M., "Electrothermomechanics of Spring-Loaded Contractile Fiber Bundles with Application to Ionic Polymeric Gel Muscles, 2nd International Conference *on* Intelligent Materials, June 94.

[13] Fessler, Mark J. G., The **ASU Finger: Design of an Anatomically-Based Finger Model for both Prosthetic and Robotic Applications,** Master's Thesis, Department of Mechanical Engineering, Arizona State University, 1990, 172pp.

[14] Delp, Scott L. and J. Petert Loan, Melissa G. Hey, Felix E. Zajac, Eric L. Topp, and Joseph M. Rosen, "An Interactive Graphics-Based Model of the Lower Extremity to Study Orthopedic Surgical Procedures," IEEE Trans. Biom. Engineering, 1990.

[15] Llorens, Will A. An Experimental Analysis of Finger Joint Stiffness. M.S. Thesis, Department of Mechanical Engineering, Louisiana State University, Baton Rouge, LA, May 1986.

5-77/46

URC97099

An Integrated Study of Tertiary Magmatism in the Rio Grande Rift Region

B.S. Penn, G. Randy Keller, and Albert Jimenez
Pan-American Center for Earth & Environmental Studies
Department of Geological Sciences
University of Texas at El Paso
El Paso, Texas 79968

INTRODUCTION

In the Mid-Tertiary, the crust of southwestern North America was extensively modified by magmatic activity. In places, the crust can be thought of as almost boiling. The magmatic activity during this time resulted from large-scale aesthenospheric upwelling of magma in a complicated back arc setting. Further upwelling occurred in the Late-Tertiary. The primary surficial manifestation of this upwelling is the Rio Grande rift which extends from Northern Colorado southward into Northern Mexico. From a physiographic viewpoint, the Rio Grande rift is an extensional feature following the crest of an uplift on which the Southern Rocky Mountains are situated. In the Rio Grande rift, distinct crustal thinning (about 5 km relative to adjacent areas) is documented from Albuquerque, NM southward. The thinnest crust found to date is located west of El Paso, TX and is approximately 28 km thick. In contrast with East Africa, the crustal thinning in the Rio Grande rift is gradual from the rift valley floor to the rift valley shoulders, perhaps reflecting the pre-rifting thermal regime.

Situated on the western shoulder of the Rio Grande rift are several large-scale (on the order of tens of kilometers across) volcanic and magmatic centers including the Datil-Mogollon volcanic field in southwestern New Mexico, the Jemez caldera in north-central New Mexico, the San Juan volcanic field in south-central Colorado, and the Princeton Batholith in central Colorado. These centers are anomalous not only because of their proximity to the Rio Grande rift but because they exhibit unusually low gravity signatures suggesting a large volumetric emplacement of silicic material in the crust. For example, geophysical studies by the UTEP group shows that the crust beneath the Datil-Mogollon volcanic field has been modified extensively and appears to contain a batholith comprising approximately 20% of the crustal material.

Thus, the rift and volcanic centers share much in terms of their tectonic evolution. This evolution has effected the lithosphere pervasively, and the associated modification of the crust has been complex. Tectonism and magmatism in this region are intimately connected and can not be studied separately. An integrated geophysical

and geological modeling approach is demanded. Such an approach must include gravity, magnetic, seismic, petrologic, topographic, and satellite data. An integrated approach is used here to ascertain the extent and magnitude of surficial and crustal aspects of Tertiary magmatism in proximity to the Rio Grande rift.

SILICIC MAGMATISM AND VOLCANISM

Several major volcanic fields dominated by silicic magmatism were formed during Mid-Tertiary time. Because these volcanic fields are well exposed, remote sensing plays a significant role in discerning and mapping the extent of the surficial aspects of these magmatic and volcanic centers. Mid-to Late-Tertiary volcanism in the southern Rocky Mountains extended from northern Colorado, into New Mexico, and merged southward with the Sierra Madre volcanic field in Northern Mexico (Figure 1). Commonly termed the "ignimbrite flareup," much of this volcanism occurred during middle to late-Oligocene times when large volumes of rhyolite to quartz latite ash-flow tuffs erupted. The talc-alkaline lavas and pyroclastic ejects from this period are believed to have merged together and covered most of central and southwestern Colorado (Lipman et al., 1970; Steven, 1975) The bulk of the volcanism associated with the ignimbrite flareup is separated into a number of individual volcanic fields (39-mile, Latir, San Juan, Questa, Raton Clayton, Jemez; Lipman, 1996).

RIFT-RELATED BASALTIC MAGMATISM

The formation of the Rio Grande rift is viewed as having post-dated Mid-Tertiary magmatism. In fact, the formation of the Rio Grande rift and the later stages of the ignimbrite flareup were synchronous (Figure 2). Subsequent magmatism near the axis of the Rio Grande rift was basaltic. Magmas during this period were clearly rift-related, volumetrically small, and mostly mantle-derived with little crustal modification. These facts agree with geophysical measurements. This basaltic-phase of magmatism is characterized by the Jemez volcanic field, the Ocate volcanic field, Raton-Clayton volcanic field, Taos Plateau volcanic field, and Springerville volcanic field. Graphing the occurrence of Mid-Tertiary magmatic events temporally suggests a rough NW to SE progression of magmatism and volcanism in Mid-Tertiary time. Compositionally, pre-Jemez magmatism and volcanism was largely intermediate talc-alkaline rocks. Late-Tertiary magmatism in the southern Rocky Mountains was dominated by hi-modal compositions of basalt and high-silica rhyolites (Christiansen and Lipman, 1972). Almost all of the high-silica and granite centers during late-Oligocene and Miocene times occurred west of the Rio Grande rift. The voluminous Questa Granites of New Mexico are the only exception.

CURRENT EFFORTS

The Datil-Mogollon volcanic field has been studied extensively and as noted above is characterized by low-crustal density suggesting the emplacement of a batholith of largely silicic material representing about 20°/0 of the total crustal volume. The Jemez caldera is currently the focus of the JTEX study. JTEX is a cooperative geophysical study being conducted jointly by University Texas at El Paso, Purdue University, and Los Alamos National Laboratory. Satellite imagery is being used to help constrain the extent of the caldera, its magmatism and rectify topographic information with subsurface measurements. By combining digital topographic and potential field data, correlations between surface and subsurface features not previously apparent are clarified. Future efforts will include similar studies of the San Juan volcanic field and the Mount Princeton Batholith.

CONCLUSION

Prior to the initiation of rifling in Southern Colorado and Northern Mexico volcanism was largely dominated by talc-alkaline felsic compositions during Mid-Tertiary time. Post-rift magmatism was composed of bimodal suites of mafic and silicic compositions. Magmas from area of the rift axis were dominantly basaltic and lacked any significant crustal signature suggesting they were derived directly from the upwelling aesthenosphere. Obviously, tectonism and magmatism combined to significantly alter the composition of the crust in this region.

Several large silicic centers of magmatic activity located adjacent to western side of the Rio Grande rift developed and appear from geophysical measurements to have altered and contributed a significant component to the crust. These centers are being investigated by integrating heretofore separate geophysical, topographic, and geologic database in an attempt to correlate surficial manifestations with subsurface measurements.

REFERENCES

Christiansen, R. L., and Lipman, P. W., Cenozoic volcanism and plate tectonic evolution of the western United States. Part 2, Late Cenozoic: *Royal Society of London Philosophical Transactions, Series* A *271:249-284,* 1972.

Lipman, P. W., Southern Rocky Mountain Volcanism, Tertiary Structural Trends, and Inception of the Rio Grande Rift in Geol. Sot. Am. Meeting, Denver, Abstracts with Programs, 28:7, p. A-270, 1996.

Lipman, P. W., Steven, T. A., and Mehnert, H. H, Volcanic History of the San Juan Mountain, Colorado, as indicated by Potassium-Argon Dating. *Geol. Soc. Am. Bull.* 81:2329-2352, 1970.

Mutschler, F. E., Larson, E. E., and Bruce, R. M., Laramide and younger magmatism in Colorado-New petrologic and tectonic variations on old themes. *CSM Quarterly 82(4):1-47,* 1988.

Steven, T. A., Middle Tertiary volcanic field in the southern Rocky Mountains, in Cenozoic History of the southern Rocky Mountains. *Geol.* Sot. *Am. Mem.* 144, Curtis, B. F. (cd.), p.75-94, 1975.

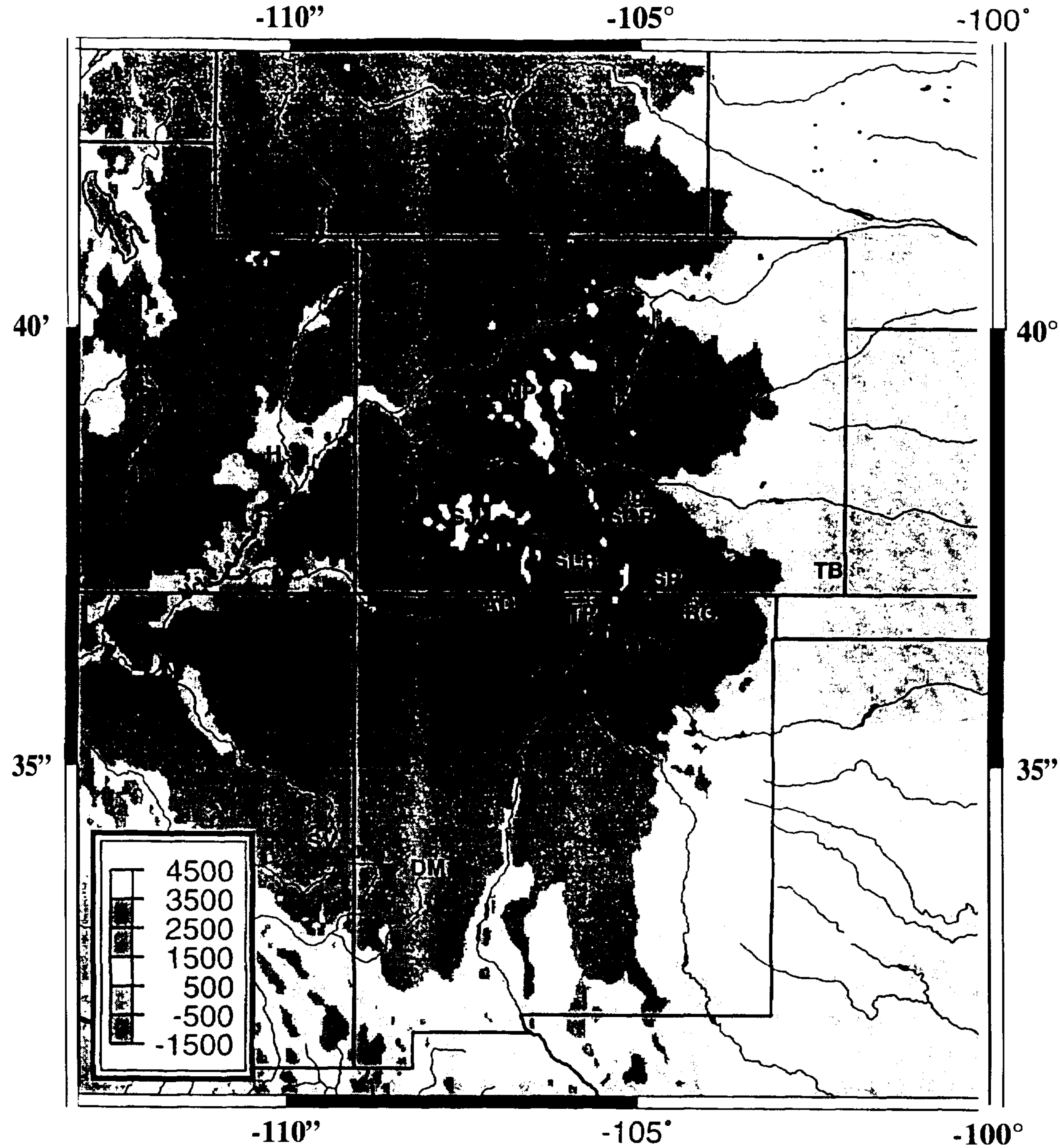


Figure 1. Locations of some of the major eruptive centers during Mid-Tertiary time. (TB: Two Buttes; SP: Spanish Peaks, MP: Mount Princeton; SJ: San Juan volcanic field, TP: Taos Plateau volcanic field, L: Latir volcanic field; AD: Archeleta-Dulce, 39:39-mile volcanic field, DM: Datil-Mogollon volcanic field, SV: Springerville volcanic field, J: Jemez mountains, RC: Raton-Clayton volcanic field, O: Ocate volcanic field, SCR: Silver Cliff-Rosita volcanic field, SLH: San Luis Hills, H: Henry Mountains.)

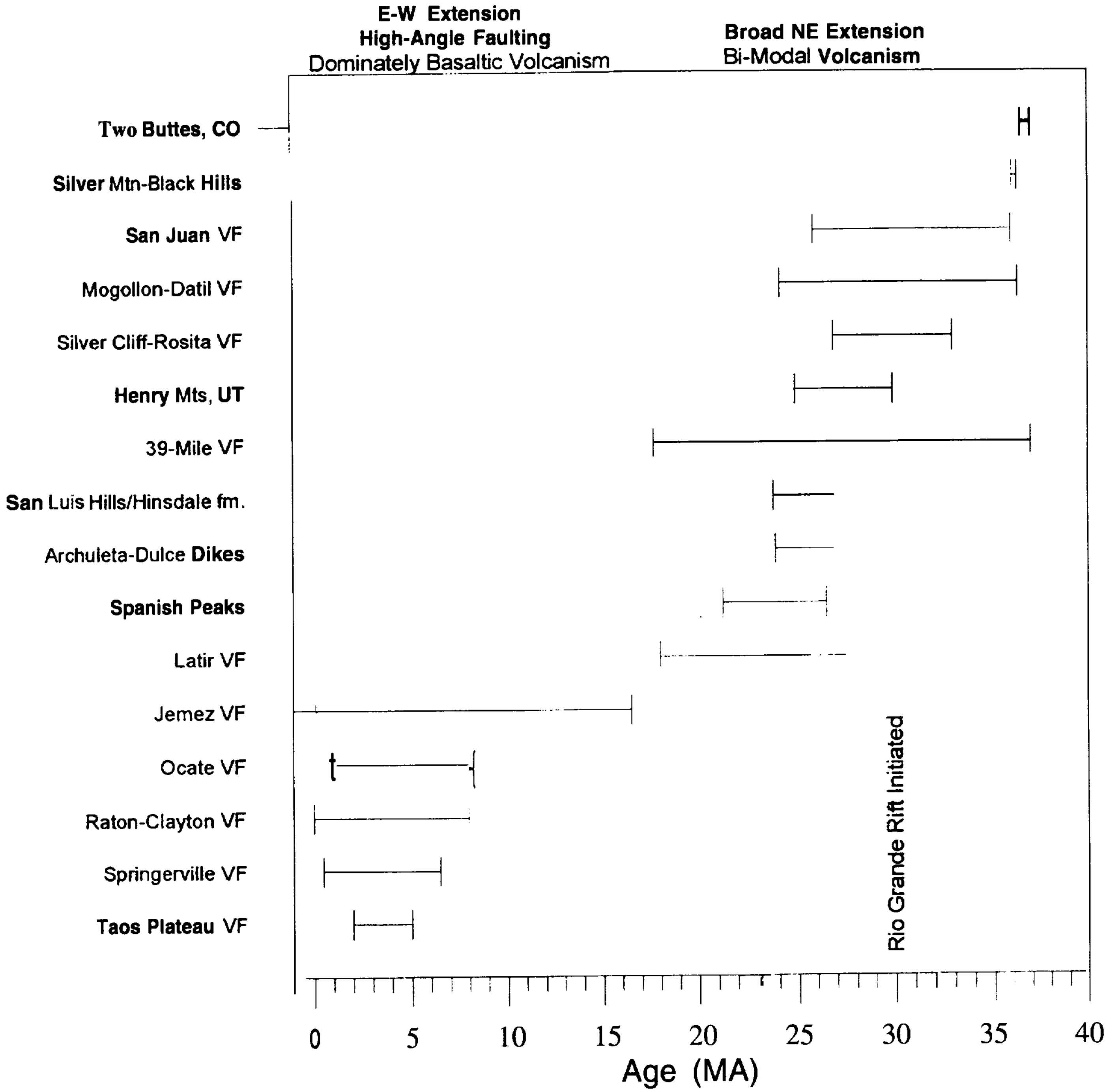


Figure 2. Occurrence of major Mid-Tertiary magmatic and volcanic centers proximal to Rio Grande rift.

URC97100

Hardware Implementation of Membership Function-Look-up Table Approach

Tim Peterson, Djuro Zrilic, Jaime Ramirez[1]*, Bo Yuan*
NASA Center for Autonomous Control Engineering
Department of Engineering
New Mexico Highlands University
Las Vegas, NM 87701

EXTENDED ABSTRACT

Time critical applications require high processing speeds and fast reconfiguration of membership function parameters. Often these requirements present a dichotomy to designers and trade off results in severe consequences to the applied solution. In addition, the behavioral complexity of control systems determines the characteristic of the controller. To cope with a wide range of applications where processing speed is of prime interest, parallel architecture of the inference processor is the most direct solution.

In this paper, considering that the primary aim of hardware implementation is the realization of high speed processing, we introduce a look-up table approach, which can realize high speed fuzzy-logic inference. Figure 1 shows a block diagram of hardware implementation of the membership function generator (MFG). Although this architecture can be extended to a multiple input multiple output generator, our primary interest is the realization of the circuit in Figure 1.

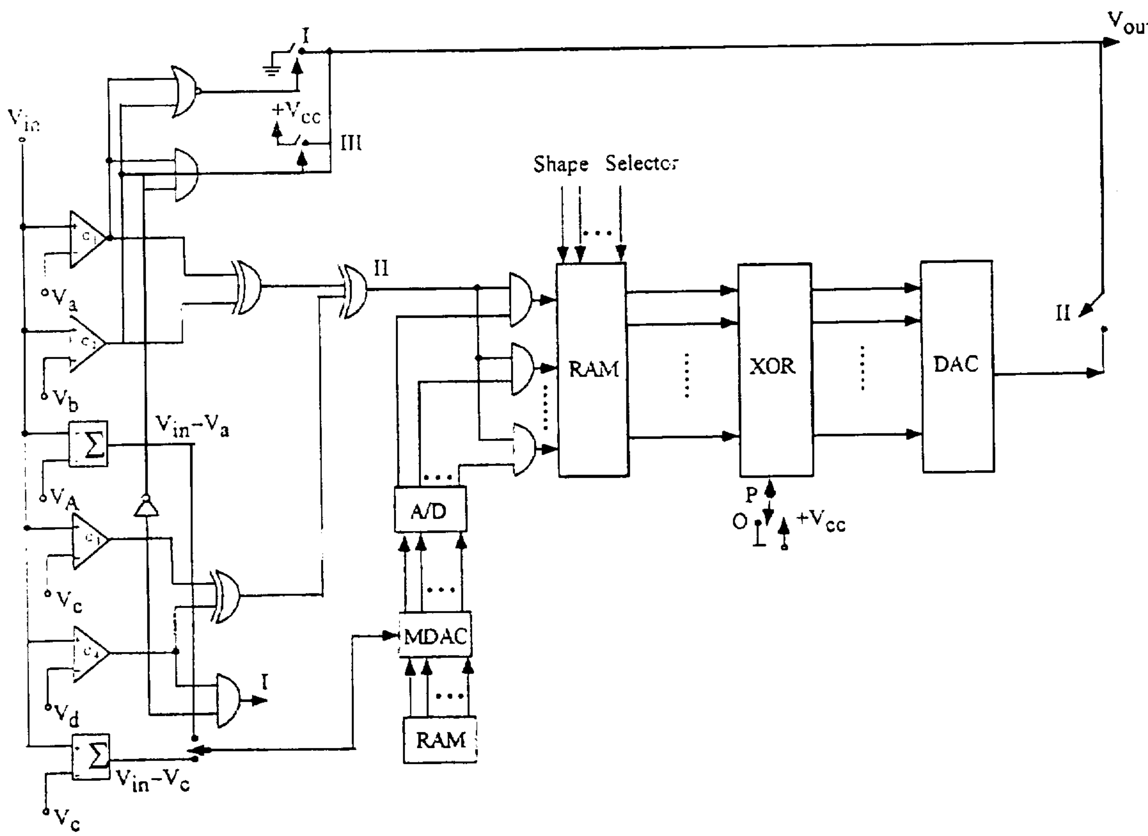


Figure 1. Block Diagram of the Proposed MFG

[1] Klipsch School of Engineering, New Mexico State University, Las Cruces, NM 88003

The look-up table approach is shape independent and can be used to realize generalized membership function as shown in Figure 2.

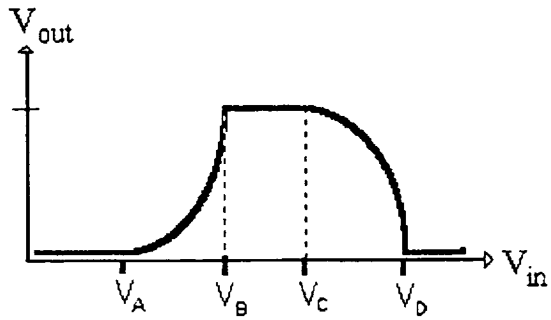


Figure 2. Generalized Shape to be generated by the circuit in Figure 1

With the stipulation that $V_a < V_b < V_c < V_d$, the general functions to be realized are J(v) and $r(v)$. The left function is

$$l(v) = f\left(\frac{Vin - Va}{Vb - Va}\right).$$

The right function is

$$r(v) = g\left(\frac{Vd - Vin}{Vd - Vc}\right)$$

The system in Figure 1 consist mainly of three parts; processor, scalar and shape generator. The role of processor is in horizontal shift along the v axis. A brief description is as follows: when $V_{in} < V_a$ outputs of comparators Cl, C_2, C_3, and C_4 are all zero, thus switch I is closed and V_{out} = O. When $V_b > V_{in} > V_a$, then switch II is closed and the contents of the shaping RAM is being addressed by scaled version of voltage difference $V_{in} - V_a$. When $V_c > V_{in} > V_b$, then switch III is closed and V_{cc} is generated at the out put. In analog fashion, because of simplicity it is possible to describe the operation of lower part of the circuit. In this case voltage difference $V_{in} - V_c$ is being processed to address content of Shaping RAM.

The role of the scalar is to generate scaled version of $V_{in} - V_a$ or $V_{in} - V_c$. Scaled number values of $1/(V_b - V_c)$ and $1/(V_d - V_c)$ are stored in a RAM. Multiplying is done by multiplying DAC. The shape generator consists of a programmable look-up table (RAM), address lines, inclusion/exclusion logic and DAC. Several shapes are stored in the RAM where the address line are used to choose the shape of a particular slope. The output of the RAM is a digital word which is the degree of belonging of the input, referred to here as the fuzzy byte. The inclusion/exclusion logic is an XOR gate array that when invoked takes the ones complement of the fuzzy byte from the RAM, essentially inverting the shape when desired. The DAC converts the fuzzy byte to an analog voltage to be processed by the rest of the fuzzy inference engine.

Details of the circuit realization of the system shown in Figure 1 using Pspice simulations are also described in the paper.

5/01/27

URC97101

Molecular Control of Polymeric Chains to Obtain a Geometrically Favorable Polyimide for Compressive Strength Studies

Odessa N. Petzold, Issifu I. Harruna*, Kofi B. Beta

Department of Chemistry and High Performance Polymers and Composites Center, Clark Atlanta University, Atlanta, GA 30314.

INTRODUCTION

Rodlike polyimides which display heterocyclic ring linkages are among the high performance polymers developed over the last two decades. The aromatic polycondensates such as, polyamides are useful candidates for high temperature and high load bearing applications.[1] As a result, there is an increasing interest in such high performance polymeric materials due to the requirements of the automotive and aerospace industries.

Aromatic polyamides have exceptional chemical resistance, high thermal and thermooxidative stability. They exhibit excellent tensile and impact strength.[1] Processing of these polymers is generally quite difficult thus limiting their utility. These limitations are due to the strong enthalpic interactions and chain rigidity exhibited by polyamides, which result in their inherent insolubility and infusibility.[2] in addition to poor processability, polyamides exhibit poor compressive strength when fabricated into films and fibers. The poor compressive strength of the films and fibers of high performance polymers is the cause of performance deficiencies encountered in their applications.

The primary approach employed to decrease interchain interactions or to reduce polymer chain stiffness involves a variety of structural modifications. This includes the introduction of bulky substituents, flexible alkyl side chains, noncoplanar moieties, and kinked and double kinked comonomers into the polyimide backbones However, such modifications generally result in the trading of high performance properties for increase processability of polyamides.

Holiday and White[4] showed that the compressive strength of polymers is dependent on secondary valence forces. To this end, rigid rod multidimensional polymers have been synthesized by Dotrong *et al.*[5] and Polk *et al.*[6,7] These approaches resulted in planar conformations and a mixture of three-dimensional and planar polymer systems, respectively.

In an effort to address the issues of processsability and poor compressive properties, we sought to prepare processable three-dimensional polyamides as excellent candidates for improved compressive strength in films and fibers. Our approach to improve processability led to the syntheses of polyamides containing the hexafluoroisopropyl group and the bicyclic ring (bicyclo[2.2.2] oct-7-ene-2,3,5,6-tetracarboxylicdianhydride) to increase volubility, while maintaining thermal stability and liquid crystallinity in polymers.

Star-like (three-dimensional) polyamides were prepared based on the 2,7-diamino-9,9-bis(4-amino-phenyl)fluorene star-like central unit. A study of the bond angles of 2,7-diamino-9,9-bis(4-aminophenyl) fluorene shows that the 9,9-phenyl substituents are at 11 3", while the bond angles between the phenyl units and the five membered ring of the fluorene molecule are approximately 107°. Therefore, the geometry at the C_9 position of the fluorene unit is almost tetrahedral (sp^3) confirming a three-dimensional fluorene based molecule as the central unit orientating the polymer chains to give the star-like polyamides.

EXPERIMENTAL

The Star-like Fluorene Unit

Synthesis of 2, 7-diamino-9, 9-bis(4-aminophenyl) fluorene The st w-like fluorene unit, 2,7. diamino-9,9-bis(4-aminophenyl)fluorene was prepared by refluxing 2,7-dinitro-9-fluorenone (7.40×10^{1} mmol) with aniline (2.44×10^{3} mmol) in the presence of aniline hydrochloride (2.22×10^{2} mmol) for 24 h under nitrogen atmosphere. The nitro functionalities were reduced by refluxing in ethanol in the presence of hydrazine monhydrate with 10% Pal/C under nitrogen atmosphere for 24 h. A precipitate was collected by filtration, washed and dried in a vacuum oven to yield 2,7-diamino-9,9-bis(4-aminophenyl)fluorene (80.59 % yield), m.p. 316-317 °C (lit.' 319.5 'C). Infrared spectrum (KBr, cm-l) 3426, 3338, 3217 (1° N-H); 3057 (aromatic C-H); 1625, 1306, 1203, 1176 (aromatic N-H); 1518, 1471, 1438 (aromatic C-C); 821 (*para* substitution); 683 (trisubstitution). Proton NMR spectrum (DMSO-d_6) 54.88, 4.92 (2H, NH_2); 6.79 (4H); 7.27 (4H); 6.47 ppm (6H). ^{13}C NMR spectrum (DMSO-d_6) 564.31 (9,9-disubstituted C); 114.03, 120.59, 123.16, 123.54, 128.13, 129.77, 143.28, 148.00, 148.06, 155.15 (aromatic carbon atoms). Anal. Calcd for $C_{25}H_{22}N_4$: C, 79.34; H, 5.86; N, 14.80. Found: C, 79.07; H, 5.94; N, 14.67.

Polymer Syntheses

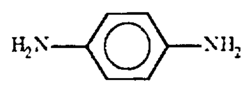
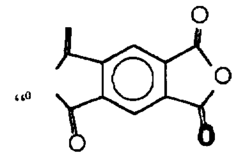
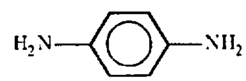
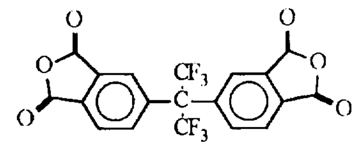
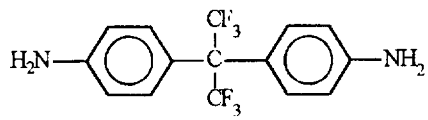
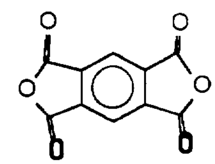
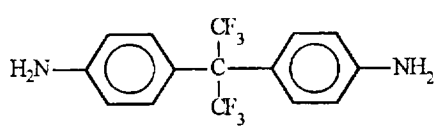
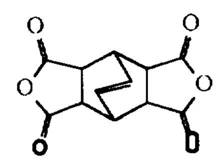
The synthesis of the linear and star-like polyamides utilized a modification of the condensation polymerization method described by Becker and Schmidt.[3] Structural representations of the primary diamines and dianhydrides used in the typical syntheses of the linear and star-like polyamides are presented in Table 1.

Synthesis of the ***Linear Polyimide*** Typically, a thoroughly dried 500 mL three-necked round bottom flask equipped with a mechanical stirrer, a nitrogen inlet tube, a reflux condenser with a drying tube, a thermometer and a bubbler was flushed with nitrogen. Primary diamine (16.2 mmol) and dry LiCl were dissolved in anhydrous NMP (100 mL) in the round bottom flask with stirring while maintaining a positive nitrogen flow. A solution of the dianhydride (17.82 mmol) in an equal volume of anhydrous NMP was rapidly injected into the reaction flask. The mixture was stirred at O °C for 1 h, and slowly heated to 30-35 "C for an additional 24 h, and finally heated to 190 °C for 24 h. Approximately 30 % of the polymer solution (60 mL) was removed, precipitated in water, washed two times with ethanol and extracted with methanol in a Soxhlet apparatus for 48 h. After drying at 60 "C for 24 h, the linear polyimide was obtained.

Synthesis of the Star-like Polyimide Typically, 2, 7-diamino-9, 9-bi s(4-aminophenyl)fluorene (2. 7 mmol) in anhydrous NMP (110 mL) was added to the remaining polymer solution of the linear polyimide and stirred at 190 °C for 72 h. The polymer solution was added to water to give a precipitate. The precipitate was washed two times with ethanol and extracted with methanol in a Soxhlet apparatus for 48 h. After drying at 60 "C for 24 h, the star-like polyimide was obtained.

Wide angle X-ray diffraction (WAXD) measurements were made using a Rigaku RU200 X-ray diffractometer equipped with a Satton camera. Nickel filtered CuKα radiation was used at 50 Kv/170 nA. Intensities were extracted from the X-ray films using a Joyce-Loeble Scandig 3 densitometer.

Table 1. Monomers Used in the Syntheses of Linear and Star-like Polyamides

POLYIMIDES	DIAMINE	DIANHYDRIDE
Linear I Star-like II		
Linear III Star-like IV		
Linear V Star-like VI		
Linear VII Star-like VIII		

RESULTS AND DISCUSSION

The structures of the linear and the corresponding star-like polyamides were confirmed via elemental analysis, FTIR, proton and ^{13}C NMR spectroscopy. The characteristic imide absorption at 1782, 1729 and 1374 cm-1 were observed in the infrared spectra of the linear and star-like polyamides. The chemical shifts observed between δ 164.63 and 175.37 ppm in the ^{13}C NMR spectra of the linear and star-like polymers correspond to the carbon atom of the imide carbonyl. ^{13}C NMR spectra of the star-like polyamides exhibit an additional chemical shift at δ 64 ppm corresponding to the quaternary carbon of the fluorene based central unit, Figure 1.

Thermal characterizations of the polymers were determined by differential scanning calorimetry (D SC) and thermogravimetric analysis (TGA) in air and under an argon atmospheres, The DSC curves of the linear polyamides I, *III*, *V* and star-like polyamides *II*, *IV*, *VI* revealed glass transition temperatures from 266 to 389 °C in air and under an argon atmospheres. Under an argon atmosphere, the DSC curve of linear polyimide *VII* exhibits an exothermic transition at 396 "C, corresponding to the enthalpy of recrystallization, and an onset of decomposition at 401 °C. In air, the DSC curve of *VII* indicates an exothermic transition at 392 "C, corresponding to the enthalpy of recrystallization, and a decomposition transition at 398 'C. The DSC curve of the corresponding star-like polyimide *VIII* exhibits an onset of decomposition at 401 "C in air and under an argon atmosphere.

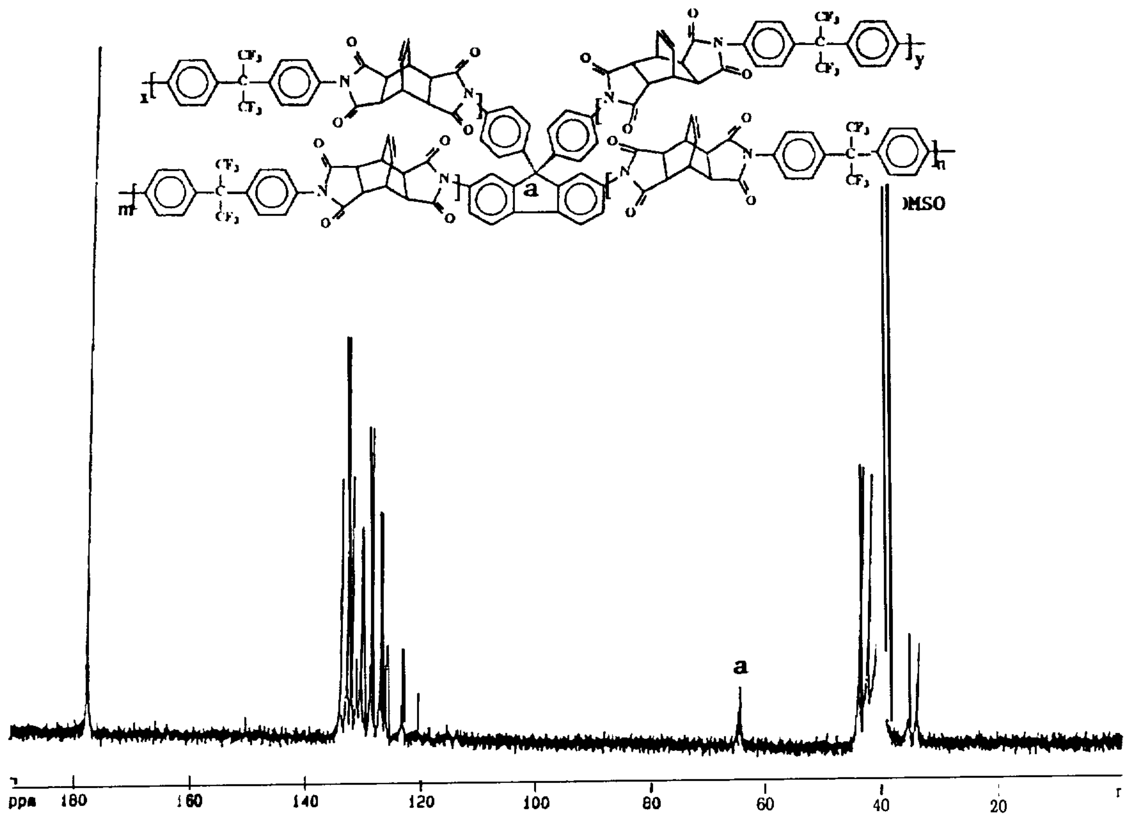


Figure 1. ^{13}C NMR spectrum of star-like polyimide *VIII* in dimethyl sulfoxide-d_6

The TGA curves of the linear and the corresponding star-like polyamides indicated thermal stability ranging from 400 to 508 °C in air and argon atmospheres. The temperatures correlating to the onset of decomposition, 5 % and 10 % weight loss of the linear and star-like polyamides under argon and air atmospheres are presented in Table 2. In general, the linear polyimides are more thermally stable than the corresponding star-like polyamides.

Table 2. Decomposition Temperatures of the Linear and Star-Like Polyamides

Polyimide	Decomposition temperature °C					
	onset		5 % wt. loss		10 % wt. loss	
	argon	air	argon	air	argon	air
linear *I*	405	381	494	443	572	505
star-like *II*	357	331	427	422	505	478
linear *III*	454	452	508	505	534	529
star-like *IV*	338	349	401	411	477	496
linear *V*	400	392	481	451	529	510
star-like *VI*	354	330	427	403	491	467
linear *VII*	357	352	414	400	430	419
star-like *VIII*	363	354	416	405	427	424

The introduction of hexafluoro groups and the bicyclic ring resulted in an improvement in the volubility of the linear and star-like polymeric systems containing these features. The polyamides *III, IV, V* and *VI*, containing the hexafluoro group and polyamides *VII* and *VIII*, containing the bicyclic ring and hexafluoro groups exhibit increased volubility in comparison to the polyamides *I* and *II* which do not contain these fictional groups. The polyamides, which displayed an improvement in volubility, were found to be soluble at 270 **w/v** concentration in organic solvents and in strong acids at room temperature, Table 3. In general, the star-like polyamides were found to exhibit greater volubility relative to the corresponding linear polyamides suggesting that the star-like polymer systems are not crosslinked.

The inherent viscosities are relatively low indicating that the polyamides are of low molecular weight. However, the viscosity of the star-like polyamides were found to be approximately twice that of the corresponding linear polyamides.

Table 3. Solubility Properties and Inherent Viscosity Measurements of the Linear and Star-like Polyamides

Polyimide	η_{inh}. g/dL	TMU	THF	DMF	DMAc	DMSO	NMP	H_2SO_4	MSA	*m*-C	Py	TCE
linear *I*	0.12'	- -	- -	- -	- -	- -	- -	+ +	+ -	- -	- -	- -
star-like *II*	0.18'	+ -	+ -	+ -	+ -	+ +	+ -	+ +	+ +	+ -	+ -	+ -
linear *III*	0.11^b	+ +	+ -	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ -
star-like IV	0.22^b	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ -
linear V	0.17^b	+ +	+ -	+ +	+ +	+ +	+ +	+ +	+ +	+ -	+ +	+ -
star-like *VI*	0.21^b	+ +	+ -	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +	+ +
linear *VII*	0.16^b	+ -	- -	+ -	+ -	+ -	+ -	+ +	+ -	- -	+ -	+ -
star-like *VIII*	0.24^b	+ -	+ -	+ -	+ -	+ +	+ -	+ +	+ +	+ +	+ +	+ -

++= soluble ; - + = partially soluble ; -- = insoluble

Volubility measurements at 2 % w/v concentrations at room temperature, 'inherent viscosity measured in concentrated sulfuric acid at 30 °C; [b]inherent viscosity measured in dimethyl sulfoxide at 30 "C; TMU, tetramethylurea; THF, tetrahydrofuran; DMF, dimethylformamide; DMAc, dimethylacetamide; DMSO, dimethyl sulfoxide; NMP, N-methyl-2-pyrrolidinone; MSA, methane sulfonic acid; *m*-C, *m*-cresol; Py, pyridine; TCE, *1*, 1,2,2-tetrachloroethane.

The X-ray diffraction scans of the linear polyimide *I* and its multidimensional analogue, polyimide *II*, exhibit semicrystalline diffraction peaks, while all the other polymers show a broad halo, indicative of amorphous nature. In the X-ray scan for linear polyimide *I*, there are four prominent peaks, centered at 4° (12.7 A), 14" (6.24 Å) and two centered at 20", with d-spacing of 4.62 and 4.20 Å, respectively. The crystal structure for this polymer has not yet been determined, but the peak at 4°, with a d-spacing of 12.7 Å, can likely be attributed to the layer lines corresponding to the ordering of the chemical repeat units along the chain axis. It is noted that star-like polyimide *II* display d-spacings identical to that of linear polyimide *I*. Efforts to determine the crystal packing of these polymers are underway.

Page intentionally left blank

URC97102

Application of Genetic Algorithms to Optimize Power Flow on a Radial Transmission Line Using Reactive Compensation

Carelle Pierre* clpierre @ neat.edu

Mohiuddin Ahmed** ahmed @neat.edu

A. Homaifar*** homaifar@ncat.edu

G. L. Lebby*** lebby@ncat.edu

Autonomous Control Engineering Center
Department of Electrical Engineering
North Carolina A&T State University
Greensboro, North Carolina 27411

Abstract

This paper presents a Genetic Algorithm (GA) based methodology for optimal capacitor placement for a general radial distribution circuit. The solution approach can optimally determine (I) the locations to install (or replace, or remove) capacitors and (II) the types and sizes of capacitors to be installed (or replaced), such that a desired objective function is minimized while the voltage constraints are satisfied. A ten-bus radial distribution feeder was used to show the effectiveness of the proposed method.

Keywords: capacitor placement, genetic algorithms, power flow, power optimization, transmission line, voltage constraints

Introduction

Optimum use of transmission lines is a planning objective of utility companies, but difficult to realize due to many constraints. Transmission line impedance elements can have an effect on the voltage at all points along the line. This effect varies with line loading and length. Transmission loss increases as more power is transferred through the line and also compounded by the line sagging due to heat increase of the conductors. Furthermore, there is the problem of voltage collapse and major blackout or shutdown of the system due to overloading of the line. Voltage collapse occurs due to gradual shifting of the equilibrium point from a region of stability to an unstable region. The main reason for voltage collapse is the lack of proper reactive support at the receiving end of the transmission line.

The performance of transmission lines can be improved by reactive compensation of a shunt type. The term compensation refers to the installation of reactive power-producing devices to achieve a desired effect in the electric power system. These effects include improved stability performance and transmission line power flow, The devices in our case are connected in parallel, or shunt, at particular points on the load side of the line. Shunt capacitance supplies reactive power to counteract the inductive load and as a result, the voltage level at the load is increased and the power loss is reduced due to a decrease in current.

Optimal capacitor placement has been approached in several ways. Chiang, Wang, and Darling (1995) used the system solution algorithms and numerical techniques. One model used a fuzzy-based approach where two membership functions have been used, one for power loss and the other for voltage sensitivity (Chin and Lin, 1995). Another model proposed to minimize

* M. S.E.E. student
** Ph.D. candidate
*** Associate professors

an objective function using simulated annealing and a greedy search technique and the problem is formulated as a combinatorial optimization problem with a non-differentiable objective function (Wang, et al, 1995). Other methods utilized the **2/3** rule, dynamic programming, a voltage-independent model, non-linear programming, and mathematical equations (Chin and Lin, 1994).

Methodology

The objective in capacitor placement in a distribution feeder, is to ensure that a proper voltage profile is maintained when minimizing the power loss. Planners must determine the proper size and optimal location of the capacitors such that the load is served as inexpensively as possible, while ensuring a reliable supply of power. Due to its size, the optimization problem can only be solved approximately or through heuristics approaches because of the CPU time required.

In this paper we have presented an evolutionary strategy algorithm (ie., genetic algorithms) to find the optimum size and location of the shunt capacitor of a radial distribution feeder when minimizing the power system loss and improving the voltage profile. The evolutionary strategy is based on the mechanics of biological evolution. GA is a search and optimization technique based on the Darwinian theory of natural selection and survival of the fittest. This optimization technique differs from conventional optimization techniques in several ways:

(1) GA searches the solution space from many different points simultaneously in parallel.
(2) the selection is based on optimizing a cost function.
(3) random search is employed, the new generations only concentrate on solutions that have been successful and survived.

In GA, when a parent spawns a new generation of children, only the best child survives to become parent in next the generation. This is similar to evolution of biological species where during evolution only the "best" survives from generation to generation.

Reactive power compensation can be achieved by the use of fixed or switched capacitors. If all variables are considered, the capacitor placement problem may become very complicated due to many non-linearities involved in the system. In this paper, only fixed capacitor placement is considered. To further simplify analyses the following assumptions are made: the system (a) is a balanced three phase system, (b) has time invariant loads, and (c) uses negligible line charging currents in comparison to capacitor banks in the distribution feeder.

Figure 1 shows an m-bus system where bus *i* contains a load and a shunt capacitor. The components are defined as follows:

$y_{i,i+1} = 1/(R_{i,i+1} + jX_{i,i+1})$: line admittance between buses *i* and *i+1*;
$R_{i,i+1}$ and $X_{i,i+1}$: resistance and reactance of the line;
Pi and Q_i : real and reactive power at bus *i;*
y_{ci} *:* shunt capacitor admittance at bus Z.

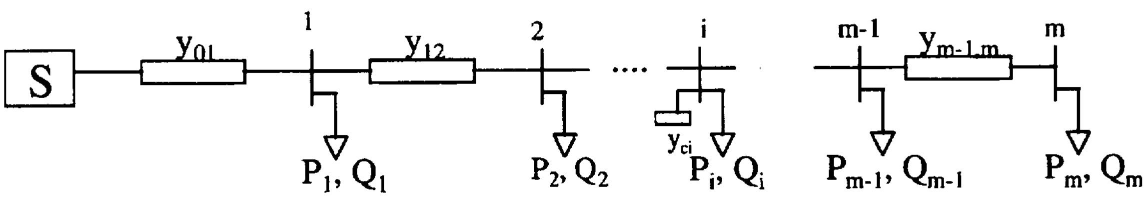


Figure 1. Single-line diagram of the distribution feeder

Our goal is to minimize the total cost of power loss in the feeder and the cost of shunt capacitors installed in the feeder. All load points are assigned some capacitor value randomly, then a power flow program is initiated to determine the new voltage profile. Next, the objective function is evaluated with the new values of the voltages to determine whether any improvement in cost is achieved in the current generation or not. This process continues until a global, or near global minimum is obtained for the cost function. The cost function consists of two terms: one is the cost of transmission line loss and the other is the cost related to the capacitor. The cost function is given as:

$$Cost = K^{P} PLoss + \sum_{j=1}^{n} K_j^C Q_j^C$$

Where

$P_{Loss} = \sum_{i=0}^{m-1} P_{loss(i,i+1)}$, K^P is the cost of power loss in a per unit quantity ($/KW/year),

and j represents the selected buses, Also, K_j^C and Q_j^C *are* the cost in dollars per kilovar and the size of the capacitor in kilovar, respectively, and are charted in Table 2 in the ***Results and Conclusions*** section of this paper. Finally,

$$P_{loss(i,i+1)} = R_{i,i+1}\left[\left|V_{i+1} - V_i\right| \cdot \left|y_{i,i+1}\right|\right]^2$$

and V_{i+1} - V_i *is* the voltage drop between the buses i and $i+1$. The admittance is defined as

$y_{i,i+1} = \dfrac{1}{R_{i,i+1} + jX_{i,i+1}}$, $R_{i,i+1}$ and $X_{i,i+1}$ are the resistances and inductances between buses i and $i+1$ respectively and are summed on the denominator of the above equation. The objective function in GA (OBJ) is then ***1/Cost.***

Procedure

In this paper, we employed GA with individuals consisting of integer-valued alleles. Every individual **in** the population is represented as a possible set of capacitors to be added to the existing bus. Therefore, individuals contain several integers, each ranging between the values of 1 (equivalent to 150kvars) and M (equivalent to 150Mkvars), where M is the maximum number of buses. The actual values that the integer values represent are multiples of 150kvars. Once GA randomly creates a population of these integer sets, or individuals, the values are analyzed to determine the power loss across the circuit. Also, we used single-point crossover with a crossover probability of 0.6, a mutation probability of 0.001, and a string length of 9 for each individual. In order to calculate power loss, line parameters such as resistance and inductane between buses, and real and reactive power at each load are needed. The power analysis routine also generates the power loss associated when the additional capacitors which are added to the circuit. GA would then create a new generation by crossing over and mutating the individuals of the old population. Again, the numbers would be analyzed for power loss and total cost. This cycle continues until an optimal, or near optimal solution is found whereby the power loss and cost is minimized subject to a specific range of voltage values across each bus.

Results and Conclusions

A test case of a ten bus radial feeder (nine load buses, shown in Figure 2) was taken for simulation and the line parameters are given in Table 1. The base values are as follows:

. base voltage= 23 kV
. base power= 13.057 MVA
. base impedance= 40.5141 **kΩ**

To obtain the per unit value of a measurement we used the equation,

$$\textit{per unit value} = \frac{\textit{actual value}}{\textit{base value}}$$

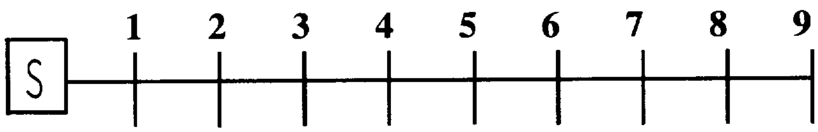


Figure 2. Radial feeder test case

Table 1

From Bus #	To Bus #	R (per unit)	X (per unit)
0	1	0.00304	0.01
1	2	0.000345	0.0149
2	3	0.0184	0.0297
3	4	0.0172	0.015
4	5	0.0489	0.0426
5	6	0.0223	0.0195
6	7	0,0507	0.0287
7	8	0.118	0.067
8	9	0.1319	0.0747

Table 2

J	1	2	3	4	5	6	7	8	9
Q^c_i(kvar)	150	300	450	600	750	900	1050	1200	1350
K^c_i($/kvar)	0.500	0.350	0.235	0.220	0.276	0.183	0.228	0.170	0.207
J	10	11	12	13	14	15	16	17	18
Q^c_i(kvar)	1500	1650	1800	1950	2100	2250	2400	2550	2700
K^c_i($/kvar)	0.201	0.193	0.187	0.183	0.180	0.195	0.174	0.188	0.170
J	19	20	21	22	23	24	25	26	27
Q^c_i(kvar)	2850	3000	3150	3300	3450	3600	3750	3900	4050
K^c_i($/kvar)	0.183	0.180	0.195	0.174	0.188	0.170	0.183	0.182	0.179

Page intentionally left blank

-/03/27

URC97103

Development of Chemically Specific Polymer Coatings for Piezoelectric Mass Sensors. A Nitric Acid Specific Sensor Based on Simple Materials.

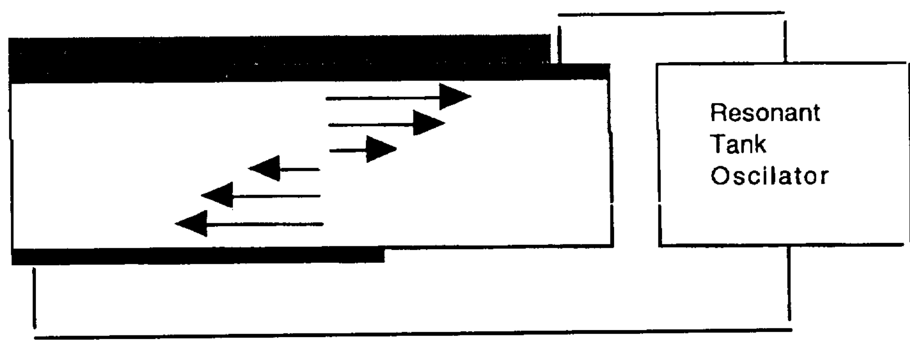
Steven Pollack and Kahsay Habte, Department of Chemistry, the Polymer Science and Engineering Program and the Center for the Study of Terrestrial and Extraterrestrial Atmospheres, Howard University, Washington, DC 20059

The determination trace gas concentration in the upper atmosphere can be accomplished either via remote sensing or through airborne instrumentation. The former situation requires a that the molecule of interest possess a chromophore of sufficient molar absorptivity and a well established absorption band where there are no interferences from other atmospheric components. The latter situation calls for a robust, compact device which is chemically specific, has a high sensitivity and may need to have a high dynamic range.

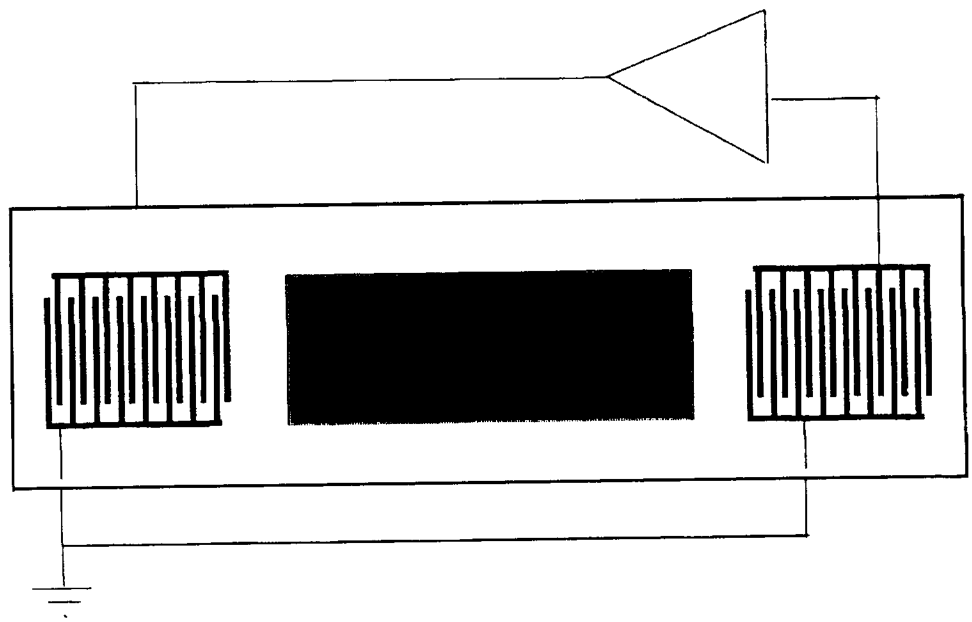
Piezoelectric mass sensors provide for a number of these latter attributes(1). These sensors rely on changes in the resonant vibrational frequency of a mechanical element in response to its effective mass. These elements can take various forms and the vibrational modes span frequencies from 8 to 400 MHz. In most cases the governing relationship which relates the mass sensing capability of these devices to their resonant frequency is the Sauerbrey equation(2)

$$\Delta m \propto {}^{\Delta f}\!/_{f_0^2}$$

where Am is the change in mass of the sensor, f_0 ***is*** the resonant frequency of the oscillator and Δf is the observed change in frequency. Two type of resonators are currently under study in our laboratories. The first is based on the bulk shear mode oscillation of a thin AT-cut quartz crystal.



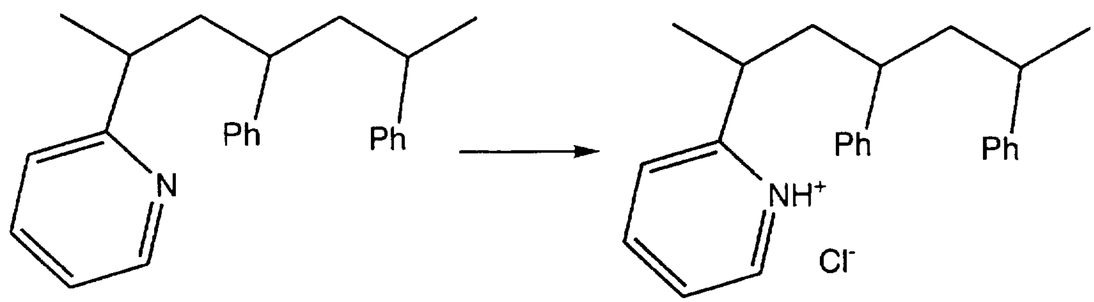
The electrodes (black) are typically gold and are applied on opposing faces of the crystal. This is commonly referred to as a quartz crystal microbalance (QCM). The first harmonic of the typical QCM is at 10 MHz and based on the physical constants of the materials it has a $\Delta f/\Delta m$ sensitivity of 1 rig/HZ. To create specificity , a suitable coating is applied (gray area). An altemative resonator is the surface acoustic wave sensor or SAW.



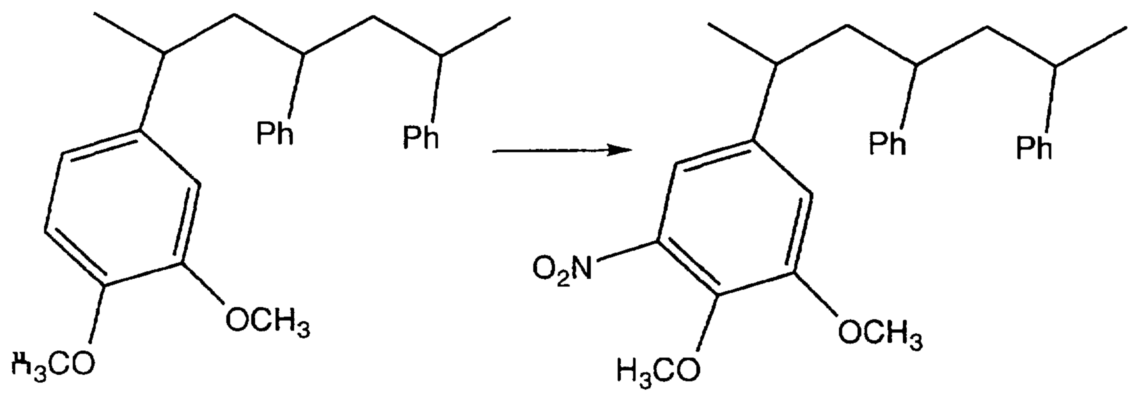
Here a slab of piezoelectric material has a pair of interdigitated electrode (IDEs) applied to opposite ends. A surface localized vibration induces an alternating current voltage in the right-hand IDE. This is amplified and fed back to the left hand IDE which in turn stimulates this vibration in the surface, traveling to the right-hand IDE. These type of devices are also referred to as analog delay lines. Here the frequency of this surface acoustic vibration is on the order of 300 MHz with the potential for a 1000-fold increase in sensitivity over the QCM. Again, a coating (gray area) is needed to make the sensor chemically specific.

There has been a great deal of effort undertaken to find coatings which are selective for a variety of gases(3, 4). Coatings can be either low molar mass materials or polymeric in nature. The latter materials have the advantage of forming thin robust films. Several polymer coatings had been proposed. For example, the co-polymer of styrene and 2-vinyl pyridine had been utilized as a

sensor for hydrochloric acid vapor. Here the specificity relies on the acid-base interaction of the gas and the pyridine moiety, the styrene acting as a diluent and for its film forming properties.



However, it was noted that in use, other acidic gases, most notably, nitric acid also gave a positive response. Additionally, it was observed that the adsorption of HCl was reversible, while that of nitric acid was not. Given that observation, we postulated that the nitric acid was reacting with the polymer, not through acid-base chemistry, but via an aromatic electrophilic substitution route. In light of this, we developed coatings which would exhibit an enhanced propensity to this nitration reaction, those based on electron-rich benzene rings, but not containing the basic pyridine moiety. In this way, we would provide a sensor for nitric acid that would not respond to hydrochloric acid. Our initial polymer was based on 3,4-dimethoxystyrene as a copolymer with styrene (PS/PDMS) (again for film forming properties control).



Polymers (of 50:50, 75:25 and 25:75 molar percent concentration) were prepared via conventional free-radical polymerization in toluene using azobisisobutyronitrile as initiator. The polymers were precipitated with methanol and dried in vacuo. Coating was accomplished via a spray deposition from toluene solutions using the change in frequency of the QCM as a monitor of coating thickness. After coating, the sensors were attached to oscillator circuits and placed in a flow cell. The PS/PDMS coatings exhibitted no mass uptake when exposed to ozone, sulfur dioxide, carbon dioxide or hydrogen chloride vapors. When exposed to nitric acid vapor, the film registered an increase in mass. We also prepared a polystyrene homopolymer film coated sensor as a negative control. We assumed that no response would be observed. Instead, we observed a positive response for nitric acid in this system. After several experiments, we concluded that the styrene moiety was sufficiently electron rich to react with the nitric acid vapor allowing for nitration. We are now exploring the sensitivity of this sensor as well as confirming the nitration mechanism as the mode of reaction. We are also investigating the

potential use of this sensor for other applications where the detection of nitric acid vapor could be important. A first application is in monitoring HNO_3 release for determining the degradation of nitrocellulose based films and movies in archival storage.

References

1. A. Mandelis, C. Christofides, ***Physics,*** *Chemisry* ***and Technology of*** *Solid* ***State Gas Sensor Devices*** (John Wiley & Sons, Inc., New York, 1993).
2. G. ***Z.*** Sauerbrey, ***Z.*** *Phys.* ***155, 206*** (1959).
3. G. Harsányi, ***Polymer*** *Films* ***in Sensor Applications*** (Technomic Publishing Co., Inc., Lancaster, PA, 1995).
4. G. G. Guibalt, J. M. Jordan, ***CRC Critical Reviews in Analytical Chemistry 19, 1-27 (1988).***

5/24/27

URC97104

Investigations on nanocrystalline thin films of CdSe for potential sensor device applications

Brajesh K. Rai[1], R.S. Katiyar[1], M.T. S. Nair[2], P.K. Nair[2], and A. Mannivannan[3]
[1]Department of Physics, University of Puerto Rico, San Juan, PR 00931-3343
[2]Lab. De Energia Solar, IIM, U. National Autonoma de Mexico, 62580 Temixco, Morelos, Mexico
[3]Dept. of Chemical Engg. And Mat. Science, U. of Minnesota, IT, Minneapolis, MN 55455-0132

Both as-deposited (AD) and after-annealing (AA) films, prepared using chemical deposition (CD) technique, have been studied by a combination of spectroscopic and structure determining techniques. AFM analysis of both AD- and AA-CdSe samples shows a wide distribution of grain size in the AD samples which becomes narrower after annealing, while the X-ray diffraction studies reveal an improvement in the crystallinity of the annealed samples. Photoluminescence spectra of the AD samples shows a weak but broad band at -2.2 (Strongly Confined Band) and -1.7 eV (Weakly Confined Band), due to the size distribution of constituent nano-clusters. Raman spectra of AD-CdSe showed, in addition to the LO and TO modes of CdSe, a surface-optic (SO) mode at - 250 cm-'; which originates due to the molecular-like small nano-clusters and disappears after annealing. Creation of surface states, leading to a maximum red-shift of the PL band to 1.68 eV, has been observed with high excitation laser powers.

During the last ten years thin films of CdSe have attracted much attention for the great fundamental[1-3], experimental[4-5], and applied[6-8] interests which they present. Bulk CdSe has a band gap of 1.73 eV. Blue-shift of the band-gap, with decrease in grain size, in the nano-crystalline phase of CdSe, has been extensively studied. Large variation of the band-gap of this material in the visible region make it a potential applicant in the preparation of solar cells and other visible light sensors. In addition, creation of localized surface states, as observed in our experiment-leading to a red-shift of the band-gap-can find applications in near-infra-red sensing devices.

The films deposited on glass, were prepared using chemical bath deposition (CBD) technique as described elsewhere[6]. Annealing of the AD films was done at 400°C in air. XRD measurements were carried our using D5000 X-ray diffractometer from Siemens Co. with CuK_{α} (1.54 A°) X-ray radiation. Nanoscope 111 AFM from Digital Instruments, Santa Barbara, CA was used for all AFM measurements. Raman and PL measurements were performed using Jobin-Yvon T64000 Spectrophotometer with subtractive pre-monochromators coupled to the third spectograph/monochromator with 1800 grooves/mm grating. Coherent Innova 99 CW Ar^{+} laser was used to provide 514.5 and 488.0 nm excitation wavelength. The normal laser power used in the experiments was about 2 m W to minimize heating and consequent changes in the crystalline state, amorphization, or oxidation of the sample. Both Raman and PL measurements were made in the -180° backscattering geometry. Data analyses were done using the software packages; namely Peak fit (Jandel Scientific Software), Auxum 4.1 (TriMetrix inc.), and Origin 4.1 (Microcal Software, Inc.).

AFM analysis of both AD- and AA-CdSe samples showed a wide distribution of grain size (2- 100 nm) in the AD samples which become narrower (2-25 nm) after annealing, while the X-ray diffraction studies reveal an improvement in the crystallinity of the annealed samples,

Fig. 1A and 1B show the Raman spectra of AD and AA samples, respectively. A LO mode at -205 cm-', whose halfwidth decreases in the AA samples due to the improved crystallinity after annealing, is observed in all the six samples. An interesting phenomenon of the disappearance of ~250 cm-' surface optic (SO) mode after annealing is seen. The presence of SO mode in AD samples reflects the molecular-like behavior of small amorphous nano-clusters.

Effect of confinement of quantum dots (QDs) on exciton has been studied both theoretically and experimentally[1-5]. The exciton spectrum of QDs depends mainly on the

quantum size effect on the conduction and valence states. Two limiting cases according to the ratio of the size of the nanocrystal, R, to the effective Bohr radius, a_b^* (56 A^0 for CdSe), of the exciton in the bulk material has been discussed by Efros and Efros.[1] In the limit R<< a_b^*, the strong confinement regime, the size quantization energy of the electron and the hole greatly exceeds the Coulomb interaction energy. Thus, in strong confinement regime the confinement of electron and hole can be treated separately with the Coulomb interaction between them as a perturbation. While, in the opposite limit R>> a_b^*, the weak confinement regime, the translational motion of the exciton is confined. Both strong and weak confinements result in a separation of the exciton energy levels, thereby, leading to a blue-shift of the PL band. AFM results of both AD and AA samples show a wide distribution in the size of QDs (2-100 rim), which encompasses both strong and weak confinement regimes as described above. Thus, in order to correlate our FL results with the theory we will consider both, strong as well as weak, confinement regimes. Energy of the exciton, in the weak confinement regime is known to show following behavior. [2]

$$E(R) = E_{bulk} + \frac{\hbar^2 \pi^2}{2MR^2} \qquad (1)$$

Where M = me + mh is the total mass of the exciton. While in the strong confinement regime, the dependence of exciton energy on the QD size is given as follows.[3]

$$E(R) = E_{bulk} + \frac{\hbar^2 \pi^2}{2R^2}\left[\frac{1}{m_e} + \frac{1}{m_h}\right] - \frac{1.8e^2}{\varepsilon R} \qquad (2)$$

Where 'ε' is the dielectric constant of CdSe. From eq^n (1) and (2), it is seen that the exciton energy, in both regimes, is proportional to 1 /R^2. However, the energy change with varying nano-cluster size is more pronounced in the strong confinement regime as compared to the weak confinement regime due to the low mass, me, of the carrier, which lies in the denominator of eq^n (2). Thus, any small variation of mean size of the QD in the strong confinement regime will be strongly reflected in terms of a large shift of PL-SCB. In the light of above discussed theoretical model we will discuss our PL results.

Fig. 2A shows the PL spectra of three AD samples in the energy range of 1.35 to 2.35 eV. The broad band centered at -1.73 eV arises due to. large nano-clusters and shows a weak quantum confinement effect. Thus, we name this band as weakly confined band (WCB). The origin of high energy band at 2.3, 2.1, and 2.0 eV in AD-4h, AD-8h, and AD- 16h samples respective] y, can be reasonably associated with the small nano-clusters; and thus, the high energy band shows a strong quantum confinement effect, accordingly we assign this band as strongly confined band (SCB). For all three AD samples, the peak position (v), intensity (I), and halfwidth ($\Delta v_{1/2}$), of both bands is shown in Table 1. Weak intensity of the two bands reveals amorphous character of the QDs, while large halfwidth of the bands indicate a polydispersity among the constituent QDs. The results of calculation of the mean radius and size-dispersion of the QDs using eq^n (2) are shown in Table I.

Following are the conclusions drawn about the three AD-CdSe samples: (A). The mean size of the QDs in the strong confinement regime increases with the increasing deposition time of the film, while the mean size of the QDs in the weak confinement regime is independent of the deposition time of the sample. Also, the rate of increase of mean QD size in the strong confinement regime decreases with increasing deposition time. (B). Dispersion in the QD size increases with increasing deposition time. The phenomenon of increase of QD size and dispersion with increasing deposition time is explained by Hodes et al.[4] as a function of Cd^{2+} and HSe in the deposition solution.

PL spectra of three AA samples are shown in Fig. 2B; Annealing results in the following changes in the PL spectra: (a) The intensity of WCB, which is very weak in AD samples,

increases significantly for all three samples with the WCB of AA-8h being most intense. In addition, WCB becomes asymmetric with its tail extending toward the high energy side. (b) As a result of annealing, the SCB decreases in intensity and shows a high energy shift in the AA-8h and AA- 16h samples (not shown). (c) A weak band on the low energy wing of intense WCB appears (not shown).

As a result of annealing, crystalline quality of bigger QDs improves, however, the smaller QDs don't seem to show any improvement in their crystalline quality. The decrease in the intensity as well as asymmetry of SCB in the AA samples can be explained as follows. During the annealing, some grains of CdSe acquire sufficient energy to dissociate themselves from very small and very large unstable nano-clusters to form stable medium-sized nano-clusters, thereby making both type of nano-clusters smaller. Thus, the annealing process results in the narrowing and shifting of the grain size distribution of CdSe toward the lower size. In the course of annealing some of the smaller nanocluster completely merge to the bigger ones; and thus a reduction in the density and size of small-sized nanoclusters (QDs) is encountered. Hence, the decrease in the intensity of the SCB in AA samples occurs due to a reduction in the density of the small-sized QDs. On the other hand, the asymmetry in the WCB is introduced as a result of a decrease in the mean size of the QDs in the weak confinement regime.

The band, on the low energy tail of WCB, at 1.41 eV has been associated with the presence of Cl and Se vacancies (V_{Se}) by Garcia-Jimenez et al.[7] The appearance of this band due to Cl impurity, in our case, is ruled out; since the ESCA studies of all of the films does not show any trace of Cl on the film surface. On the other hand, the annealing of AD films can result in the creation of Se vacancies. Thus, we associate the band at 1.41 eV to the presence of Se vacancy.

Fig. 2C shows the effect of different excitation laser power on the PL spectra of AA- 16h sample. With increasing laser power, the WCB is observed to show a red-shift in its peak position and its intensity decreases. The PL spectra of AA4h, AA8h, and all three AD samples (not shown) exhibit a similar dependence of excitation power. It is known that annealing of the AD-CdSe films done at temperatures greater than 400°C results in a degradation of the film quality. [9] Exposure of the film spots to high laser powers, for which the spot temperature goes beyond 400°C, leads to a degradation of the nanocrystals contained within the spot and consequent] y the PL intensity decreases. Thus, with increasing excitation laser power, which causes an increasing degradation of the film's crystal linity, a decrease in the PL intensity is observed.

With increasing laser power the WCB shows a red-shift in its peak position and with an excitation laser power of 12m W the peak of WCB of AA 16h shifts to 1.68 eV. This phenomenon suggests the creation of surface states in the forbidden gap. Existence of localized surface states in the forbidden gap of nano-clusters has been discussed by Brus.[3] More recently, an analogous situation in the nano-crystalline $BaTiO_3$ has been observed, [9] where the origin of surface states were attributed to Ti-O dangling bonds. With high excitation laser power, which results in a degradation of nano-crystals present in the sample and an oxidation of the surface, similar dangling bonds of Cd and Se involving oxides are expected to form which may give rise to a continuous distribution of surface states, as observed in the present experiment. However, a thorough investigation of behavior and origin surface states will require PL studies at low temperatures.

In conclusion, the CD-CdSe thin film has been found to show a quantum confinement effect. Using PL and Raman results, we have demonstrated a simultaneous existence of bulk and nano-cluster/-crystal phase in the AD as well as AA samples. The PL data, together with AFM and XRD results, give a very precise quantitative estimation of size distribution and crystallinity of the grains in the sample. With our PL experiments, performed using high excitation power, we

established that a red-shift in the band-gap (E_g <1.73 eV) can be achieved with the creation of surface states in the forbidden gap. We have shown that the spectroscopic and structure determining techniques when applied together can be a very sensitive and precise tool to observe the quantum confinement effect in the CdSe nano-crystals.

References

1. Al. L. Efros and A. L. Efros, Sov. Phys.-Semicond. 16,772 (1982).
2. Selvakumar V. Nair, Sucharita Sinha, and K.C. Rustagi, Phys. Rev. B 35,4098 (1 987).
3. L. Brus, J. Phys. them. 90,2555 (1986).
4. Gary Hodes, Ana Albu-Yaron, Franco Decker, and Paulo Motisuke, Phys. Rev. B 36,4215 (1987).
5. R. Garuthara and G. Levine. J. Appl. Phys. 80,401 (1996).
6. M.T.S. Nair, P.K. Nair, Ralph A. Zingaro, and Edward A. Meyers, J. of Appl. Phys. **74**, 1879(1993),
7. J.M. Garcia-Jimenez, G. Martinez-Montes, and R. Silva-Gonzalez, J. Electrochem. Sot. 7, 2048 (1992).
8. M. S. Shaalan and R. Muller, Solar Ceils 28, 185 (1990).
9. Meng Jinfang, Brajesh K. Rai, R. S. Katiyar, unpublished.

Table I: Spectral data for luminescence in six samples.

Sample	WCB1 ν (I, $\Delta\nu_{1/2}$)	WCB2 ν (I, $\Delta\nu_{1/2}$)	SCB1 ν (I, $\Delta\nu_{1/2}$)	R σ	SCB2 ν (I, $\Delta\nu_{1/2}$)	R σ
AD-4h	14420 (54, 2892)		18600 (134,2958)	2.57 0.86		
AD-8h	14110 (229, 1948)		16710 (332, 1844)	3.34 1.19	18470 (215, 1910)	2.61 0.56
AD-16h	13850 (135, 1567)		16230 (253, 1362)	3.67 1.16	17470 (333, 4165)	2.95 2.28
AA-4h	13940 (7276, 862)	14220 (3460, 1503)	17740 (120, -4000)			
AA-8h	13950 (17000, 834)	14250 (8170, 1445)	17730 (1 50, -4000)			
AA-16h	13720 (1 1220, 928)	14160 (4624, 1526)	17000 (90, -4000)			

ν, I, and $\Delta\nu_{1/2}$ represent the peak frequency (cm-'), the peak intensity (counts/see), and half-width, in that order. And R and σ represent mean radius (rim) and dispersion (rim) of the QD-size, respective] y.

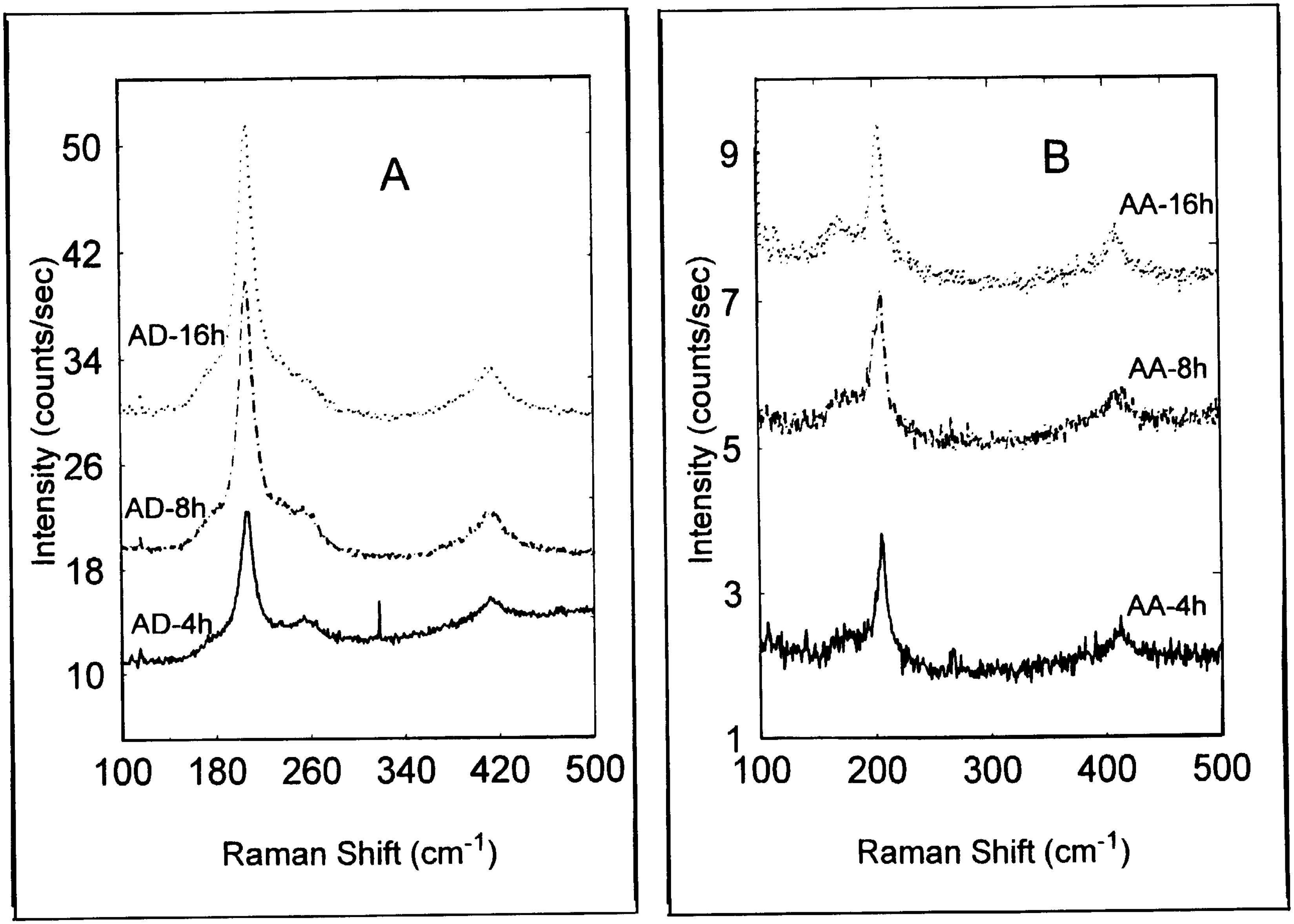


Fig. 1: Raman spectra of (A) AD and (B) AA samples. For clarity, the AD-8h and AD-16h spectra have been shifted upwards by 10 and 20 counts/sec, respectively. While the AA-8h and AA 16h spectra have been shifted upwords by 3 and 6 counts/sec, respectively.

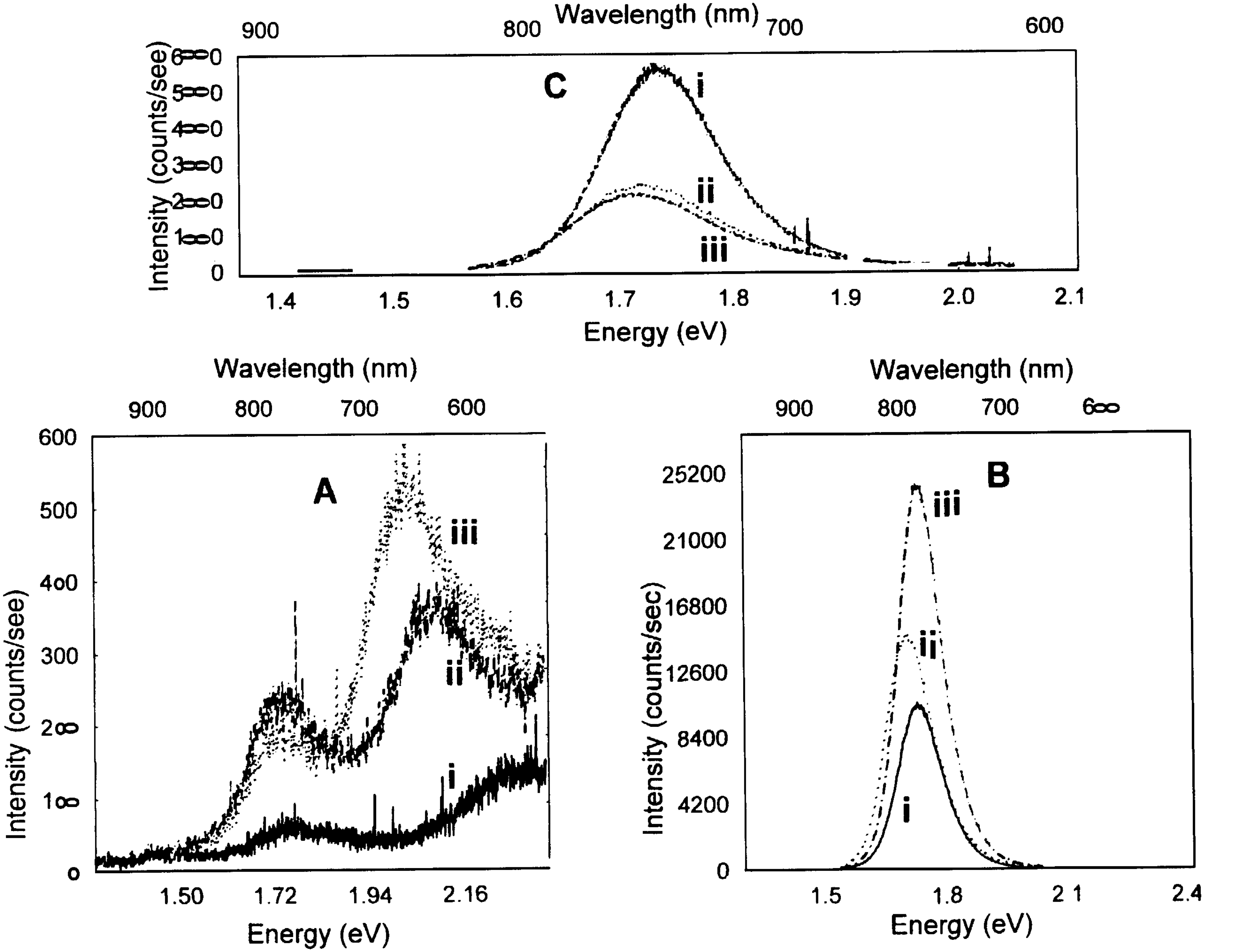


Fig. 2: PL spectra of: A) CD films deposited for i. 4h, ii. 8h, iii. 16h; B) AA films deposited for i. 4h, ii. 8h iii. 16h; C) CD-16h film with laser excitation aser powers of . 2 mW, ii. 8 mW, iii. 12 mW

URC97105

Can Chlorine Anion Catalyze the Reaction of HOCl with HCl?

S. L. Richardson, J. S. Francisco[2], A. M. Mebel[3], and K. Morokuma[3]

[1]Department of Electrical Engineering, Howard University, Washington, DC 20059;
[2]Department of Chemistry and Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana 47907-1393;
[3]C. L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322

Abstract

The reaction of $HOCl + HCl \rightarrow Cl_2 + H_2O$ in the presence of Cl^- has been studied using *ab initio* methods. This reaction has been shown to have a high activation barrier of 46.5 kcal mol^{-1}. The chlorine anion, Cl^- is found to catalyze the reaction, viz. two mechanisms. The first involves Cl^- interacting through the concerted four-center transition state of the neutral reaction. The other mechanism involves the formation of a $HCl \cdot HOCl \cdot Cl^-$ intermediate which dissociates into $Cl_2 + Cl^- + H_2O$. The steps are found to have no barriers. The overall exothermicity is 15.5 kcal mol^{-1}.

Understanding the underlying chemistry behind stratospheric ozone depletion is central to finding ways of curtailing further loss of ozone in the Antarctic Ozone Hole. One of the major reaction processes implicated in the catalytic removal of ozone is the heterogeneous reaction of HCl with $ClONO_2$ to give HNO_3 and Cl_2.[1,2,3] The Cl_2 can subsequently photolyze to produce Cl atoms which can then participate in homogeneous catalytic cycles for ozone destruction through

$$Cl + O_3 \rightarrow ClO + O_2 \quad (1)$$
$$ClO + O \rightarrow Cl + O_2 \quad (2)$$
$$\text{net:} \quad O + O_3 \rightarrow 2O_2 \quad (3)$$

Recent field measurements of HCl and ClO concentrations in the stratosphere show that the reaction of HCl with $ClONO_2$ is critical in determining the winter and spring chlorine budget

within the polar vortex. [4,5] However, extensive experimental studies suggest that the heterogeneous reaction of HCl with $ClONO_2$,

$$HCl + ClONO_2 \rightarrow HNO_3 + Cl_2 \tag{4}$$

may not occur in one step but two viz.[6,7,8]

$$ClONO_2 + H_2O \rightarrow HOCl + HNO_3 \tag{5}$$
$$HOCl + HCl + Cl_2 + H_2O \tag{6}$$

The net reaction of steps 5 and 6 is reaction 3.

A key step is the reaction involving HOCl with HCl (reaction 6). Experimentally,[9] this reaction is known to be slow in the gas-phase; but little is known about the details of the energetic that could explain the slow rate. Abbatt et al.[6,7] have asserted that reaction 6 can occur on polar stratospheric cloud (PSC) particles. However, the existence of solvated chlorine in the form of Cl^- on PSC particles has been suggested.[5,6,7] This has raised the question of whether Cl^- can catalyze the reaction of HOCl with HCl (reaction 6). In this paper we use *ab initio* calculations to ascertain if Cl^- can lower the activation barrier for the HOCl -t HCl reaction.

The geometries of reactants, products and transition states have been optimized using the *ab initio* MP2 method.[10] Two basis sets have been used for optimization: 6-3 1G(d) and 6-31 1++G(2d,2p) which includes diffuses- and p-functions and an extra set of d-functions on all heavy atoms. The diffuse functions are added because we believe these to be essential for accurate description of geometries and energetic for anions. Single point energies have been calculated using the CCSD(T) method with the extend 6-31 1++G(3df,3pd) basis set.

Geometries of the transition state for the reaction of $HOCl + HCl \rightarrow Cl_2 + H_2O$ is shown in Figure la. The number without the asterisk are geometric parameters that have been calculated at the MP2/6-3 1G(d) level, and the numbers with the asterisk are at the MP2/6-311++G(2d,2p). The transition state is a four-center reaction. The HCl approaches HOCl orienting the hydrogen

of the HCl toward the oxygen of HOCl. The HCl bond length increases from the length of the HCl bond in isolated hydrogen chloride. The ClO bond length also increases from that in HOCl. The transition state is a non-planar structure. The four atoms involved in the four-center ring is out-of-plane by 5.9°. Moreover, the hydrogen on HOCl is out of the plane of the ClOH' group by 98°. The transition state is a true first-order saddle point, with one negative frequency.[11] The relative energetic are given in Table 1. The experimental heat of reaction[12] at OK for the HOCl + HCl reaction is -18.0 kcal mol^{-1}. At the CCSD(T)/6-311++G(3df,3pd) level we predict a value of -15.4 kcal mol^{-1}, which is in reasonable agreement with the experimental results. The calculated barrier of the same level of theory is 46.5 kcal mol^{-1}. Such a high barrier suggest that the homogeneous reaction will be very slow which is consistent with the experimental observations.

In the presence of Cl-, the anion prefers completing through the hydrogen not involved in the four-center ring. The anion has a dramatic effect on the transition state. The transition state along with structural parameters are given in Figure lb. The transition state is also a first-order saddle point with one imaginary frequency.[13] Comparing the HCl bond, this bond is longer in the transition state involving the anion than without the anion present. The ClO bond in the transition is significantly elongated as a result of the Cl-. It is also interesting to note that the atoms in the four center ring is more out-of-plane than in the neutral reaction; 24.0° versus 5.9°, respectively. These changes in the transition state for HOCl + HCl involving Cl^- compared to without the Cl^- indicate that the transition state occurs earlier in the entrance channel. The structural changes in the transition state also suggest that breaking the ClO bond in HOCl and the HCl bond in HCl in the transition state is easier by the presence of the Cl-. The activation barrier for HOCl + HCl + Cl" formation is 8.3 kcal mol^{-1} at the CCSD(T)/6-311++G(3df,3pd) level. The activation barrier for the neutral reaction (3) is found to be very high and the reaction is slow. However,

involvement of Cl^- anion dramatically changes the potential energy surface. These calculations shows that the involvement of the anion lowers the neutral barrier by 38.2 kcal mol^{-1}. Moreover, these calculations suggest that the anion could play a significant role in reducing the energetic of the HOCl + HCl reaction.

While exploring the potential energy surface of the HOCl + HCl reaction in the presence of Cl-, we found an alternate mechanism for the formation of the $Cl_2 + H_2O$ products. A stepwise mechanism involving the formation of anion complexes such as $Cl^- \cdot HOCl$ and $Cl^- \cdot H_2O$ was found. The $Cl^- \cdot HOCl$ anionic complex reacts with HCl in a barrier-less process to form a HCl - HOCl . Cl^- intermediate. This intermediate is shown in Figure 2. At the MP2/6-31G(d) level, the ClO bond in the $HCl \cdot HOCl \cdot Cl^-$ intermediate is elongated nearly 0.6 Å from the ClO bond in HOCl, and the H'Cl bond is elongated by 0.7 Å from the uncomplexed HCl. We have performed a vibrational frequency analysis to determine if the intermediate is a saddle or minimum energy point on the potential energy surface.[14] The vibrational energy calculations reveal that the structure has all real and positive frequencies. The energy of the intermediate formation from the reaction of HCl + HOCl and Cl^- is calculated to be -37.3 kcal mol^{-1} at the MP2/6-31G(d) levels, and reduces to -38.0 kcal mol^{-1} at the CCSD(T)6-311++G(3df,3pd) level, The formation of Cl_2 + H_2O + Cl^- requires 22.5 kcal mol^{-1} at the CCSD(T) level. As a result, one can view an alternate mechanism for HCl + HOCl catalysis by Cl^- as involving a step-wise mechanism first involving the formation of a Cl^- . HOCl complex which then attaches to HCl forming an HCl . HOCl . Cl^- intermediate. The next step involves the detachment of Cl_2 and Cl^- . H_2O complex. At low temperatures this mechanism involving the $HCl \cdot HOCl$. Cl^- intermediate may be preferable. We should point-out that such an attachment/detachment mechanisms has been suggested for the reaction of HCl with $ClONO_2$ catalyzed by NO_3^-.

More importantly, the present calculations suggest that anions, such as Cl-, can play a major role in promoting the reaction of HOCl with HCl. The energetic effects on the HOCl + HCl reaction are considerable. Moreover, anions such as nitrate and/or sulfate which are important ions in atmospheric aerosols may play similar roles in the energetic of the HOCl with HCl reaction.

Acknowledgments

One of us (SLR) wishes to thank the Cherry L. Emerson Center for Scientific Computation at Emory University for its support and hospitality as a Emerson Fellow at Emory during the 1995-1996 academic year. One of us (JSF) would like to thank the JPL S upercomputing Project for support of this computing research. The JPL Supercomputing Project is sponsored by JPL and the NASA Office of Space Science and Application. We also wish to acknowledge support from NASA Grant NAGW-2950 for this research.

References

[1] Solomon, S.; Garcia, R. R.; Rowland, F. S.; Wuebbles, D. J. Nature, 1986,321,755.

[2] Toon, O. B.; Turco, R. P. Sci. Am. 1991,264,68.

[3] Brune, W. H.; Anderson, J. G.; Toohey, D. W. Science, 1991,252, 1260.

[4] Webster, C. R.; May, C. D.; Toohey, D. W.; Avallone, L. M.; Anderson, J. G.; Newman, P.; Lait, L.; Schoeberl, M. R.; El King, J. W.; Chan, K. R. Science, 1993,261, 1130.

[5] Haas, B. M.; Crellin, K. C.; Kuwata, K. T.; and Okumura, M. J. Phys. Chem. 1994,98,6740.

[6] Abbatt, J. P. D.; Molina, M. J. J. Phys. Chem. 1992,96,7074.

[7] Abbatt, J. P. D.; Beyer, K. D.; Fucaloro, A. F.; McMahon, J. R.; Wooldridge, P. J.; Zhang, R.; and Molina, M. J. J. Geophys. Res., 1992, 97, 15819.

[8] Chu, L. T.; Gen, M. T.; Keyser, L. F. J. Phys. Chem, 1993,97, 12798.

[9] DeMore, W. B.; Sander, S. P.; Golden, D. M.; Hampson, R. F.; Kurylo, M. J.; Howard, C. J.; Ravishankara, A. R.; Kolb, C. E.; Molina, M. J. Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling, Evaluation No. 10, NASA, Jet Propulsion Laboratory, Pasadena, CA.

[10] Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; DeFrees, D. J.; Baker, J.; Stewart, J. J. P. and Pople, J. A., GAUSSIAN 92/DFT (Gaussian, Inc. Pittsburgh, PA 1993).

[11] Vibrational frequencies at the MP2/6-31G(d) level of theory for the neutral HCl + HOCl $\rightarrow$ $Cl_2 + H_2O$ reaction are: 3604,2041, 1598,948, 811,365,318,222, 1288 i cm^{-1}.

[12] Chase, M. W., Jr.; Davies, C. A.; Downey, J. R., Jr.; Frurip, D. J.; McDonald, R. A.; and Syverud, A. N. JANAF Thermochemical Table, 3rd cd., J. Phys. Chem. Ref. Data, 1985, vol. 14.

[13] Vibrational frequencies at the MP2/6-31G(d) level of theory for the $[HCl + HOCl + Cl^-]^{\ddagger}$ transition state are: 3273, 1751, 1544, 1046, 891, 552, 352, 271, 232, 67, 49, 643 i cm-l.

[14] Vibrational frequencies at the MP2/6-31G(d) level of theory for the HCl . HOCl . Cl^- intermediate are: 3793, 2999, 1746, 959, 606, 384, 370, 321, 201, 162, 136, 39 cm-*.

[15] Mebel, A. M.; Morokuma, K. J. Phys. Chem. 1956, 100,2985.

Table 1. Relative Energetics (in kcal mol^{-1}) with Respect to HOCl + HCl Catalyzed by Cl^-

Level of Theory	$HOCl + HCl \rightarrow Cl_2 + H_2O$ $\Delta H_{r,o}$	$HOCl + HCl \rightarrow Cl_2 + H_2O$ $H^\ddagger$	$[HOCl + HCl + Cl^-]^\ddagger$ $H^\ddagger$	$Cl^- + HOCl + HCl \rightarrow HCl \cdot HOCl \cdot Cl^-$ $\Delta H^0_{r,o}$	$HCl \cdot HOCl \cdot Cl^- \rightarrow Cl_2 + H_2O + Cl^-$ $\Delta H^0_{r,o}$
MP2/6-31G(d)	-10.1	52.3	8.1	-37.3	27.2
MP2/6-311++G(2d,2p)	-16.8	49.7	10.3	-41.4	24.6
CCSD(T)/6-311++G(3df,3pd)[a]	-15.5	46.5	8.3	-38.0	22.5
Experiment	- 8 ○				

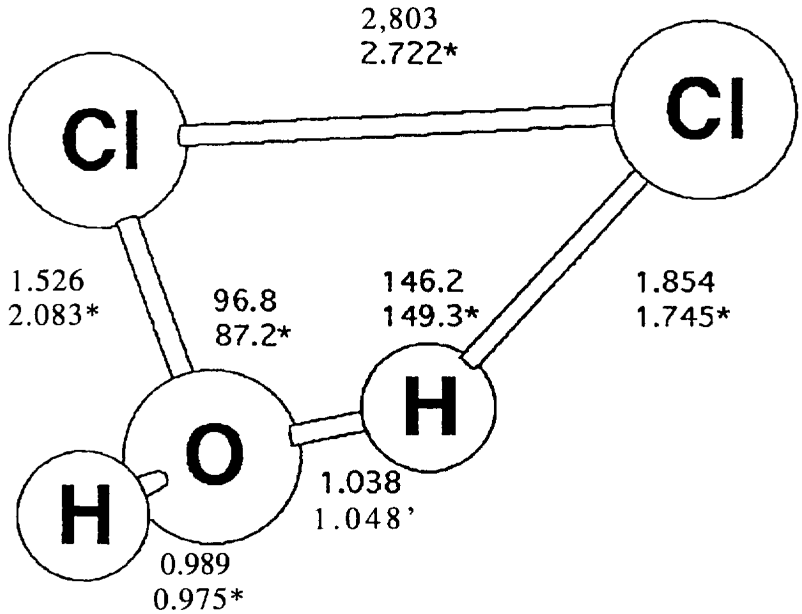
[a] Geometries from the MP2/6-311++G(2d,2p) level are used.

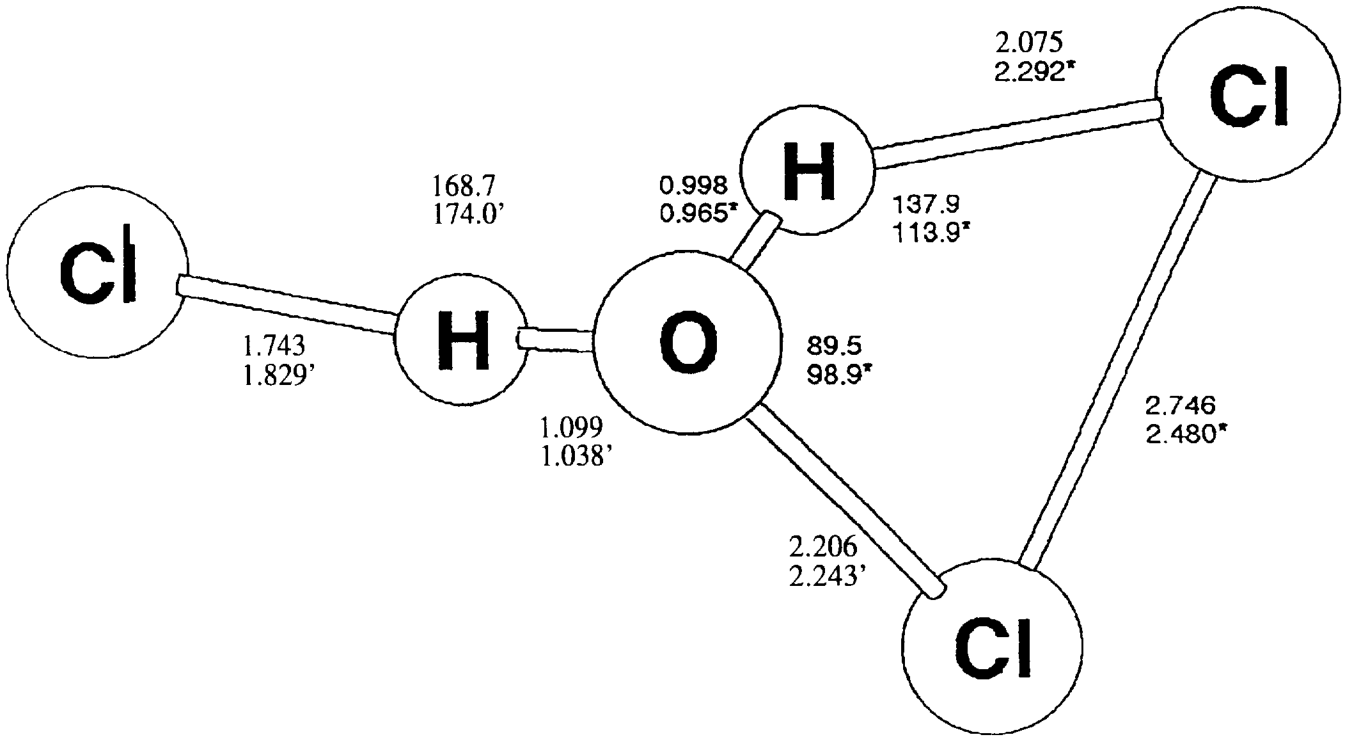
Figure Captions

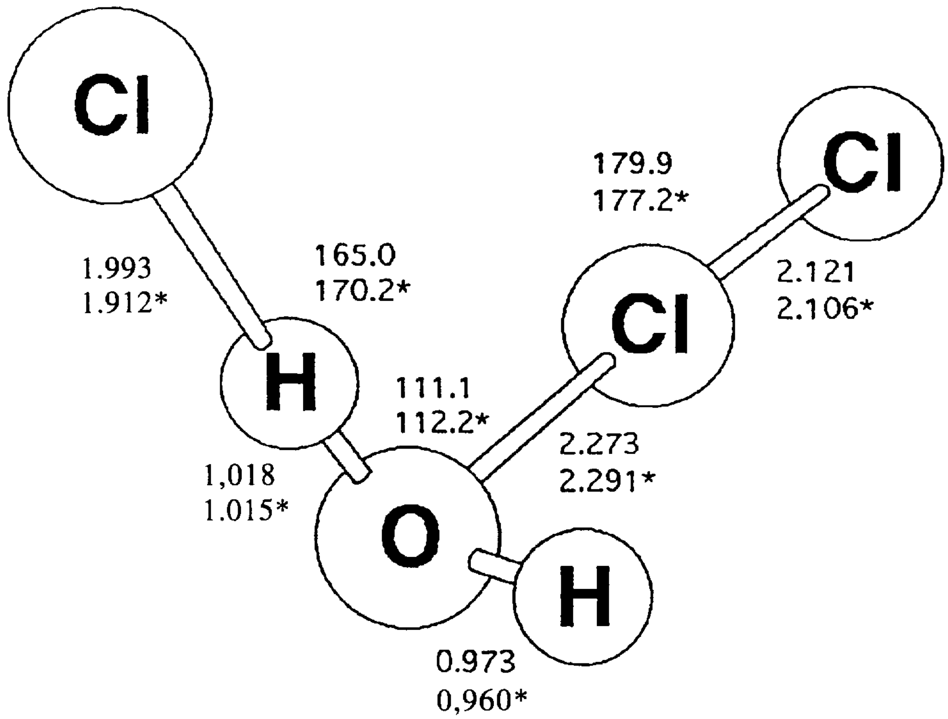
Figure 1. (a) Transition state for HCl + HOCl $\rightarrow$ Cl_2 + H_2O for the neutral reaction. Numbers without asterisk are calculated at the MP2/6-31G(d) level, numbers with the asterisk are at the MP2/6-311++G(2d,2p) level of theory.

(b) Transition state for HCl + HOCl + Cl^- + Cl_2 + H_2O + Cl-with Cl^- interacting through the four-center HCl + HOCl transition state. Numbers without asterisk are calculated at the MP2/6-31G(d) level, numbers with the asterisk are at the MP2/6-311++G(2d,2p) level of theory.

Figure 2. HCl · HOCl · Cl^- intermediate. Numbers without asterisk are calculated at the MP2/6-31G(d) level, numbers with the asterisk are at the MP2/6-311++G(2d,2p) level of theory.


2,803
2.722*
Cl
Cl
1.526
2.083*
96.8
87.2*
146.2
149.3*
1.854
1.745*
H
O
H
1.038
1.048'
0.989
0.975*


2.075
2.292*
Cl
H
168.7
174.0'
0.998
0.965*
137.9
113.9*
Cl
H
O
1.743
1.829'
89.5
98.9*
2.746
2.480*
1.099
1.038'
2.206
2.243'
Cl


Cl
1.993
1.912*
165.0
170.2*
179.9
177.2*
Cl
2.121
2.106*
Cl
H
111.1
112.2*
2.273
2.291*
1,018
1.015*
O
H
0.973
0,960*

Page intentionally left blank

/ · / 5/06/31

URC97106

Minority Universities Systems Engineering (MUSE) Program at the University of Texas at El Paso

Mary Clare Robbins, Ph. D.
Department of Mechanical and Industrial Engineering
The University of Texas at El Paso

Bryan Usevitch, Ph. D.
Department of Electrical and Computer Engineering
The University of Texas at El Paso

Scott A. Starks, Ph. D.
Director, Pan American Center for Earth and Environmental Sciences
The University of Texas at El Paso

Introduction

In 1995, The University of Texas at El Paso (UTEP) responded to the suggestion of NASA Jet Propulsion Laboratory (NASA JPL) to form a consortium comprised of California State University at Los Angeles (CSULA), North Carolina Agricultural and Technical University (NCAT), and UTEP from which developed the Minority Universities Systems Engineering (MUSE) Program. The mission of this consortium is to develop a unique position for minority universities in providing the nation's future system architects and engineers as well as enhance JPL's system design capability. The goals of this collaboration include the development of a system engineering curriculum which includes hands-on project engineering and design experiences.

UTEP is in a unique position to take full advantage of this program since UTEP has been named a Model Institution for Excellence (MIE) by the National Science Foundation. The purpose of MIE is to produce leaders in Science, Math, and Engineering. Furthermore, UTEP has also been selected as the site for two new centers including the Pan American Center for Earth and Environmental Sciences (PACES) directed by Dr. Scott Starks and the FAST Center for Structural Integrity of Aerospace Systems directed by Dr. Roberto Osegueda. The UTEP MUSE Program operates under the auspices of the PACES Center.

Needs Addressed by the Program

Academic departments within universities are extremely capable of creating curricula. However, curricula is often lacking in connectivity. Senior-level capstone design courses help tie together courses within a single discipline, but do little to provide connection between courses in various departments. As technology has grown increasingly sophisticated, it has become clear that real-world problems require teams of engineers and scientists from every discipline. The coordination of this effort, systems engineering, is rarely taught as a discipline in and of itself. It requires the focus of a project. All this contributes to a general lack of understanding of what a practicing engineer does. Students often graduate and have little grasp of the communication and teamwork skills that engineers must have to do their jobs effectively. Additionally, students are often unaware of the numerous career opportunities which abound in industries such as aerospace,

Objectives of the Program:

As a result of this project, two major objectives will have been accomplished:

Academic and Professional Development of our students:

Provide students access to career opportunities in the aerospace industry, and government research and development laboratories.

Provide students with first-hand knowledge of the engineer's workplace and overcome the problems associated with the leap that must be made from coop/summer internship programs to the first job.

Stimulate students to pursue graduate study.

Development of a system engineering curriculum:

Incorporation of systems engineering concepts in all engineering disciplines.

Regular offering of an interdisciplinary systems engineering course sequence,

Activities Associated with the Program

Fall, 1995

Student recruitment represented the first activity of the project. An informational meeting was held at UTEP in mid-November to provide students an opportunity to obtain information related to the year's activities and to meet the UTEP faculty who will be involved. Kim Leschly, a systems engineer at JPL, attended the meeting. Selection of the students was made by a committee and based on completion of an application and academic record.

Spring, 1996

Students attended videotaped lectures provided by JPL. The topics of these lectures pertained to Satellite Subsystems and included: Systems Engineering Overview, Thermal Control Design, Communications Systems Design, Aerospace Mechanisms, and Power. At the completion of the spring semester 3 undergraduate and 1 graduate student were chosen to work at JPL during the summer.

Summer. 1996

Design Activity at JPL

Three UTEP undergraduate students and one UTEP graduate student spent nine weeks at JPL where they participated in the program designed by JPL in order to become acquainted with the System Engineering Tools and Project Engineering Tools available in the JPL Project Design Center. During this session they met and worked with students from CSULA and NCAT. This opportunity provided an ideal setting for building teamwork skills and to lay the foundation for these geographically-distanced students to continue their collaborative efforts when they return to their respective schools. These students participated in all aspects of the design of the satellite which they named Urania. Their final presentation was a formal design review which included a presentation of the Satellite Implementation Plan. The Review Board was comprised of a number of JPL experts in satellite design.

Design Activity at UTEP

UTEP was the site of a design activity during the Summer of 1996 in order to verify the approach which ultimately will be used in the development and institutionalization of the desired systems engineering curriculum. Five undergraduate students worked two months at UTEP preparing a detailed satellite communications subsystem design under supervision of Dr. Bryan Usevitch. This particular subsystem was one which UTEP agreed to assume the lead of the development.

Fall, 1996

Activities in the Fall included attendance at the AIAA/USU Conference on Small Satellites. Students presented their work in Poster Sessions at the Conference. Faculty and students from all three participating schools met with Kim Leschly of JPL, and devised a plan for continuing the work on the project. An electronic newsletter was one vehicle chosen to keep members informed of the progress of other groups. Several students at UTEP are continuing the work of the group in their senior design projects.

Rationale

The proposed plan of activities presents a cost-effective approach to assist UTEP in achieving the objectives stated earlier. Specifically, designing a satellite can serve as a unifying theme for bringing together students, faculty, and staff to study concepts from systems engineering. We felt that our students would be keenly interested in the prospects of designing a system that may ultimately be flown in space, and thus the recruitment of students for the proposed project was readily accomplished. Due to the wide range of engineering disciplines that must come together to accomplish such a design, the choice of a satellite project will lead to the desired interaction among departments in UTEP's Colleges of Engineering and Science. This in turn should lead to the development of a model curriculum in systems engineering which will find an audience among students from a wide range of academic majors. The existence of the consortium will enable us to draw upon the expertise present at other schools and will also provide a wider range of student backgrounds upon which we can test the effectiveness of the curriculum to be designed.

The involvement of JPL personnel and the availability of resources at JPL is critical to the success of the project. Through this project, UTEP students will be able to meet and work on a continuing basis with professionals at JPL. This experience is especially important for many UTEP students since the majority of UTEP students come from backgrounds which have prevented them from knowing first-hand what to expect in the engineering workplace. In general, they have little experience in relating to professionals holding graduate degrees other than faculty. Furthermore, the El Paso region is not one which contains large numbers of aerospace employers. The time that UTEP students spent at JPL will give them first-hand knowledge of career opportunities that exist in the aerospace industry.

Evaluation Plan

During 1995, UTEP was awarded a major grant from the National Science Foundation called Model Institution for Excellence (MIE). This major proposal lays the programmatic and systemic reform foundation for UTEP's strategic mission to increase its production of science, engineering and mathematics degrees awarded to underrepresented minority students. A major activity associated with the preparation of this proposal was a self-study in which over 200

faculty and staff participated. The self-study resulted in the development of goals for the numbers of students to be graduated over a five-year horizon.

We like to think that the educational activities associated with the MUSE Program as comprising an important part of the overall minority recruitment and retention efforts of UTEP in line with the goals of the MIE proposal. Programs such as this one are vital to UTEP in its efforts to obtain the desired increases in underrepresented minority student degree production in science, engineering and math. Through this project, UTEP has been able to provide some stipend support for approximately 15 students at the undergraduate level, as well as for one graduate student over the next year.

In terms of collecting evaluation data, UTEP has in place two offices which will support this effort. The Center for Institutional Evaluation Research and Planning (CIERP), along with the Office of Institutional Studies (OIS), will track students who participate in the Satellite Design experience. In particular, we will be interested in determining how many of these students, by virtue of their participation in the program, will ultimately enter graduate school in Electrical, Mechanical or Civil Engineering, as well as Geological Sciences. Another goal of this project will be to stimulate students to enter careers after the bachelor's degree in the aerospace industry. We will track students after graduation to determine the numbers of students that do exactly this.

Systems Engineering Curriculum Initiatives at UTEP

A major goal of the MUSE program has been to strengthen the ability of the participating universities to educate their students in the field of systems engineering. Along these lines, UTEP has adopted a staged strategy for the introduction of systems engineering concepts and courseware in its curricula. A philosophy has been adopted that the courses to which these efforts are to be applied should be regularly scheduled courses, preferably required by major. Additionally, UTEP has selected courses as candidates for the integration of systems engineering principles which have as their goal the design and development of student projects. It is the judgment of the UTEP faculty that group-oriented design and development are the most natural venue for learning systems engineering concepts. The following sections provide a near term plan for its curriculum development activities relating to systems engineering.

Senior Design Project Course Sequence in Electrical and Computer Engineering

The Senior Design Project course sequence is a two-course sequence which is required for all students majoring in electrical and computer engineering. In this course sequence, students work in groups of 3-4 to define, design and build a prototype product. The course sequence was previously modified through the efforts of an IBM Faculty Fellow on loan to UTEP. The course was restructured to introduce students to some concepts of systems engineering: preparation of requirements, specification, and other related documents. Furthermore, students are introduced to several forms of project planning, such as PERT charting. To enhance communication skills, students are required to make numerous written and oral reports. The Department of Electrical and Computer Engineering enrolls the largest number of students of any department within the College of Engineering. The Senior Design Project course sequence has won praise for its success in preparing students for future careers.

During the Fall 1996 semester, UTEP will use the existing Senior Design Project course sequence in the Department of Electrical and Computer Engineering as the setting for its curriculum reform efforts in systems engineering. Systems engineering courseware developed at JPL will be integrated into the lecture portion of the Senior Design Project course sequence. In particular, UTEP will pursue the introduction of systems engineering elements such as design to

cost principles, project trade-offs, and other concepts which have previously not been a part of the Senior Design course sequence. UTEP also plans to pursue the possibility of involving members of JPL's technical staff in the teaching of various systems engineering topics. It is envisioned that this might be accomplished via distance learning and visits to the UTEP campus. By augmenting and enhancing the systems engineering component of this course sequence, the entire class of graduating seniors in electrical and computer engineering will be impacted. This should be on the order of 100 students per year, the bulk of whom are underrepresented minorities.

All majors in the College of Engineering at UTEP offer capstone senior design courses. Once success is shown in terms of infusing systems engineering concepts into the Senior Design Project course sequence in electrical and computer engineering, UTEP will investigate mechanisms for replicating the process with other majors in the College.

Special Topics Course in Mechanical Engineering

UTEP will offer an upper-division special topics elective in the Department of Mechanical and Industrial Engineering during the Spring 1997 Semester. The subject of the courses will be internal combustion engines and a potential project of the course will be development of a propane-powered vehicle. This elective course will be an excellent candidate for integration of systems engineering courseware and instructional materials. Systems engineering concepts such as effective team building, subsystem specification and design and others will be introduced in the course to provide a framework for the diverse student groups who will be responsible for the various vehicle subsystems. It is believed that approximately thirty upper-division engineering majors will receive instruction in systems engineering concepts and processes. Additionally, this course will provide UTEP with a multi-disciplinary setting into which systems engineering instructional materials can be placed.

Other Courses

The concepts of systems engineering will also be taught to engineering students in the Introduction to Engineering course sequence. These courses are multidisciplinary, and required for students planning to pursue an engineering degree at UTEP. An additional course, communication systems engineering, will be offered by UTEP in Fall, 1997. This course will be offered through the Electrical Engineering department, and will focus on systems design as it relates to communications.

Summary

In line with the requirements set forth by the Accreditation Board for Engineering and Technology, and in light of current trends in engineering education, the College of Engineering at UTEP has embarked upon a process that has resulted in the teaching of engineering design across the curricula. Presently, students are given opportunities to participate in engineering design activities at all stages of their development beginning with the freshman year and culminating with senior capstone design classes. Many of these activities are group-oriented. It is the expectation of UTEP that by integrating systems engineering concepts into selected courses such as those presented earlier, students will be able to master systems engineering principles within the context of group-oriented projects. Also, the approach presented above will enable UTEP faculty to evaluate the need for establishing a specific course in system engineering.

Acknowledgments

We wish to thank NASA for its support of this project through the grant NCCW-0089. Also,we wish to extend our gratitude to Al Paiz at NASA's Jet Propulsion Lab for his vision in the development of this project.

Modeling the Breakthrough of Chloride in the Vadose Zone

*ROBERTSON, G. C., S.A. ABURIME, R.W. TAYLOR, J.W. SHUFORD and T.L. COLEMAN
Center for Hydrology, Soil Climatology, and Remote Sensing (HSCaRS)
Alabama A&M University, Normal, Al. 35762

ABSTRACT

Models simulating water and solute movement in the soil are in abundance. Invariably, these models are based on some form of the convective-dispersion equation, implying that the water and solutes follow an average pathway through the soil. Chloride was leached through a Decatur silt loam (Rhodic Paledult) soil. Soil solution samples were collected from a 60cm depth in the soil. Model performance was evaluated by ploting the predicted concentration vs. observed. Overall, the model worked well in predicting the breakthrough of chloride.

INTRODUCTION

Improved model parameterization of complex flow paths depends on the ability to accurately characterize the undisturbed vadose zone. Results obtained from simulation models are heavily dependent on the input parameters which in turn are dependent on the accuracy of data obtained with a vadose zone sampling (VZS) device. Since most models fail to predict the variations of contaminant concentrations in groundwater (Ritter et al., 1987), their results should be interpreted with caution. This failure occurs because these models assume an average flow path and velocity of migration and do not consider preferential flow paths through which solutes can quickly bypass the biologically active root zone, thereby decreasing the potential for chemical degradation (Richard and Steenhuis, 1988).

Apart from preferential flow pathways (e.g., wormholes, root holes, cracks, and /or wetting front instability), chemical persistence and adsorptive properties can affect a chemical's ability to degrade to a harmless form during its residence time in the biologically active region of the vadose zone or pass through the vadose zone to ultimately contaminate groundwater. The knowledge of transport processes and soil-chemical interaction are required for adequate predictions of contaminant movement in such soils and they are also essential for characterization assessment, and cleanup of groundwater contamination. Therefore, the purpose of this study was to predict the breakthrough of chloride using the classical steady-state convection-dispersion equation and the mobile-immobile two region transport model from the Concentration Distance Time FIT program, CXTFIT, (Parker and van Genuchten, 1984).

<u>Model:</u>

The Concentration Distance (x) Time Fit program by Parker and van Genuchten, (1984), will be used in evacuating the predominant flow path in this study. Thus, more review is done on the classical convective dispersion and the mobile-immobile phase transport equation.

<u>Equation for Water Flow</u>

To predict solute transport under transient flow conditions, one must first be able to accurately describe water flow. This has been accomplished by coupling terms accounting for various forms of solute flux to the Richard's equation.

Darcy's equations was assumed to describe the flow of water in a saturated and/or unsaturated soil. The equation for flow in one dimension in an isotropic soil is

$$q = -K(\theta)\frac{\partial h}{\partial z} \qquad [1]$$

where q = volumetric soil-water flux (m/s), $K(\theta)$ = Hydraulic conductivity (m/s) which is a function of the soil water content(@) (cm^3/cm^3), h=Hydraulic head, (m) = $\Psi m + z$ (unsaturated soil), z Gravity potential (equivalent to soil depth (m) measured positive downward), and Ψm= soil-water matric potential or capillary pressure.

The volumetric soil-water flux is smaller in magnitude than the velocity of the water moving through the tortuous pore sequences. The average pore velocity (v) is obtained by dividing the soil-water flux (q) by the volumetric soil-water content (θ) Equation [1] describes a steady state soil water condition.

Theoretical Models on Chemical Movement in Heterogeneous Soils

The one dimensional flow of a non-adsorbing, chemically inert solute through a porous medium is usually analyzed with the convection-dispersion equation [2] which is based on the assumption that field soils are homogeneous:

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} - v\frac{\partial c}{\partial x} \qquad [2]$$

where c= solution concentration (ML^{-3}), D= dispersion coefficient (L^2T^{-1}), t= time (T),v= average pore-water velocity (LT^{-1}), and x= distance (L).

Equation [2] assumes steady state water flow at a constant soil-water content, and no interactions between the chemical and the solid phase. It therefore implies that all soil water participates freely in the convective and dispersive transport of chemicals.

Adsorbing Chemicals

When chemical adsorption is considered, an additional term must be added to Equation [2] to account for the interaction between the chemical and the solid phase. This is accomplished by redefining Equation [2] as

$$\frac{\partial c}{\partial t} + \frac{\rho\partial S}{\theta\partial t} = D\frac{\partial^2 c}{\partial x^2} - v\frac{\partial c}{\partial x} \qquad [3]$$

The solution of Equation [3] depends on knowledge of the relationship between the adsorbed concentration, S, and the solution concentration, c. The relation between adsorbed and solution concentrations may be described by a linear isotherm of the form

$$S = k\,c \qquad [4]$$

where k is an empirical distribution coefficient ($M^{-1}L^3$). Substitution of Equation [4] into Equation [3] gives the transport equation

$$\frac{R\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} - v\frac{\partial c}{\partial x} \qquad [5]$$

where the retardation factor, R is given by

$$R = 1 + \rho k/\theta \qquad [6]$$

where S = adsorbed concentration (MM-1),
θ = volumetric soil-water content (L^3L^{-3}) and
p = bulk density (ML^{-3}).

Chemical Breakthrough Curves (BTC)

Equation [5] is more conveniently used to analyze BT-curves by the introduction of dimensionless variables such as:

$$T = vt/L \qquad [7]$$

$$P = vL/D \qquad [8]$$

$$X = x/L \qquad [9]$$

$$C_1 = \frac{c - c_i}{c_o - c_i} \quad [10]$$

Introducing these variables into [5] gives

$$\frac{R\,\partial C_1}{\partial T} = \frac{1}{P}\frac{\partial^2 C_1}{\partial x^2} - \frac{\partial C_1}{\partial x} \quad [11]$$

where R= retardation factor,C 1= dimensionless solution concentration in the mobile region, T = number of pore volumes, P= column Peclet number and X= dimensionless distance along the flow direction.

MATERIALS AND METHODS

Description of Field Experiment

The research was performed at the Alabama A&M University (AAMU) Winfred Thomas Research Station. In 1992, eight (8) 0.2 ha experimental plots were established. The predominant soil type at the location is Decatur silt loam (Rhodic Paleudult). Each plot is spaced 20 m from each other with a lateral subsurface tile drain installed at 1 m depth. The plots are isolated from the surroundings by perimeter tiles at 1.8 m depths to intercept subsurface flow. The plots are adjacent to an underground sampling station that will not be discussed in this paper which provides an overfall where runoff measurement and sampling devices are located. Plots 1, 2, 7, and 8 will be the on] y plots discussed in this paper,

Site Chemical **and** Physical Characterization

This site was characterized by collecting soil samples (disturbed and undisturbed) randomly at different locations and depths on the plots. These samples were used to determine soil physical, chemical and hydraulic properties such as soil texture, soil porosity, volumetric water content, bulk density, and hydraulic conductivity. Standard methods as described by Black (1965) and Page (1982) were used. These data are required as input in the computer models which will be used to analyze the suitability of the sampling devices.

Schedule of Pesticide-Sampling

Data were collected more frequently during the initial period tier solute application and also after each rainfall event. The plots received natural rainfall and samples were collected with the vadose zone sampling devices (wick and gravity lysimeters) according to the following schedule: a) During or one day after precipitation (rainfall). b) three days after precipitation. c) seven days after precipitation. d) Ten days after precipitation. e) fourteen *days after precipitation.*

Tracer Analyses

Water samples were collected in pre-cleaned amber glass bottles and stored at 4°C until analyzed. The soil and water analysis were conducted at the AAMU's Environmental Research Laboratory.

Chloride was chosen to measure the water's velocity because this anions move at approximately the same rate as water through negatively charged clay soils. Chloride concentrations were determined with an Ion Chromatography. Diluted effluent solutions were analyzed for chloride by inserting a 5 ml sample into a Dionex series 200 ion chromatagraph with an IONPAC AS4 anion exchange column and a model AMMS-1 anion exchange membrane suppressor. A high purity grade helium was used as the gas carrier in this study. Data reduction was accomplished by using the accompanying AI-400 software.

Modeling

The breakthrough curves (BTCs) from each of the sampling method were simulated with the Concentration Distance (x) Time Fit program, CXTFIT, (Parker and van Genuchten, 1984). This model is used to fit velocity, dispersivity, and adsorption partition coefficients for each location sampled. By comparing the observed concentrations with the predicted concentrations at different times, the effectiveness of each vadose zone sampling (VZS) method can be determined. Furthermore, by comparing the distribution of flux and velocity, the heterogeneity within the VZS location in each of the plots can also be determined.

Breakthrough Curves (BTC) of the Lysimeters

Breakthrough curves of the applied chemical sampled with the various sampling devices are reported and discussed as follows. The combined cells chemical breakthrough curves for the wick and gravity lysimeter in plot 1 are shown in Figure L Breakthrough of chloride (0.347 g), took place 1 day after 3.3 cm of rainfall from the wick lysimeter (Fig. la) on Day 20 after chemical application (ACA).

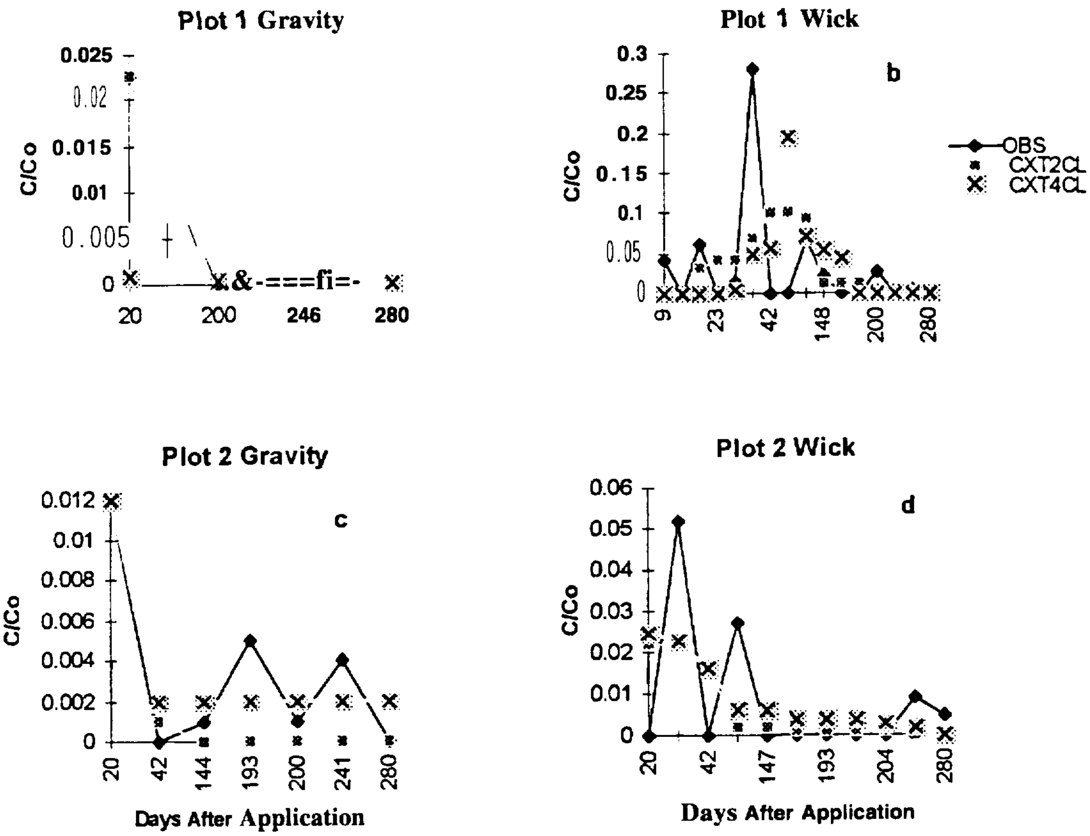


Figure 1. Predicted Chloride (Using CXTFIT Modes 2 and 4) vs. Observed breakthrough Curves for Chloride and Pesticides in Plots 1 and 2

More chloride (4.99 g) was detected on Day 23 (June 10, 1994) in response to a total rainfall of 10.16 cm on June 6 and 7. Furthermore, the 4.44 cm rainfall of Day 40-42 led to the breakthrough of 22.75 g of chloride probably due to the high antecedent soil moisture content of about 0.43 kg/kg in the plot.

In plot 1 peak concentrations of the observed for the wick lysimeter (Figure lb) arrived earlier than the predicted concentrations of modes 2 and 4. Overall, the shape of the curves indicates an under prediction of concentration by the model with model II predicting the best because of its closeness of fit.

As for gravity lysimeter in plot 1 (Figure la), the chloride breakthrough was detected about one day after the 10.16 cm of rainfall (June 6 &7, 1994). After the June 6 and 7 high rainfall events, there was not enough rainfall to cause any other breakthrough in the gravity lysimeter.

In the gravity lysimeter (Figure la), model II worked best because its predicted peaks were identical to the observed. Model I obviously missed the first initial breakthrough.

In Plot 2, the chloride in the wick lysimeter increased to 9,791.18 ug/ml on the third day following the rainfall (Day 23 ACA) from a concentration of 48.67 ug/ml measured on day 20 for the wick. Then, in response to a 5.10 cm rainfall on Day 144 there was a breakthrough of 2.21 g of chloride, Finally, on day 241 there was more chloride (0.69 g) effluent from the wick in response to the 2.79 cm rainfall that fell on day 240 ACA (Figure id).

In plot 2, peak arrival time of the observed was identical to the time of the models. However, concentrations were significantly different, with the observed having the highest concentration followed by model II and then by model I. Model 11 seem to have work the best for the model in this plot because

its peaks were closest to the observed peaks (Figure Id). The same trend was noticed in the gravity lysimeter (Figure 1c) as well, except for the initial breakthrough on day 20 which was well predicted by both modes.

In plot 7 (Figure 2b) an exponential reduction of chloride concentration in the wick lysimeter occured on day 23. In the gravity lysimeter, a four peaks were observed in responce to rainfall on days 20, 193, 200, 280, respectively (Figure 2a).

In plot 7, the model again worked well in the wick lysimeter (Figure 2b) with both modes slightly under predicting the highest observed concentration on day 23. In the gravity, the model worked well by showing a good to excellent fit of the model to the chloride data (Figure 2a).

In plot 8, the first breakthrough of chloride occurred on day 23 after a cumulative rainfall of 16.7 cm in depth. This was followed by a higher breakthrough on day 144.

The model seem to work well for the wick lysimeter (Figure 2d) except for on day 23 and day 144 were it under predicted the observed concentration. As for the gravity lysimeter (Figure 2c), the model poorly fits the data from day 200 through 280, as a result, no trends were noticed from the predictions.

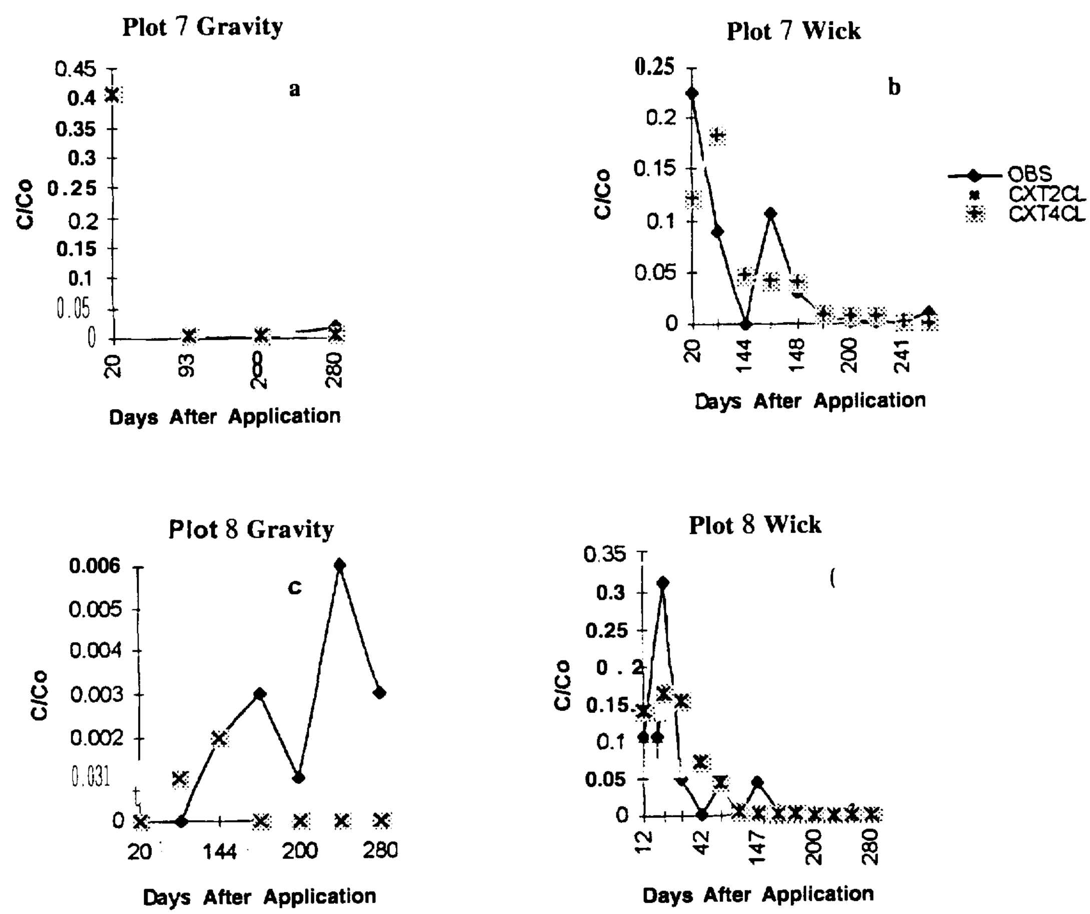


Figure 2. Predicted Chloride (Using CXTFIT Modes 2 and 4) vs. Observed breakthrough Curves for Chloride and Pesticides in Plots 7 and 8

CONCLUSIONS

The models used to predict the breakthrough curve of chloride seemed to work well by producing good to excellent fit curves when plotted against the observed for most of the plots. However, model II seem to have worked better than model I in predicting breakthrough of chloride. Model II also seemed to have worked best in predicting matrix and perferential flow observed in the lysimeters, while model I best predicted perferential flow which was observed in the gravity lysimeters.

REFERENCES

Parker, J.C. and M.Th. van Genuchten. 1984. Determining Transport Parameters from Laboratory and Field Tracer Experiments. Virginia Agricultural Experiment Station Bulletin 84-3.

Parlange, J.Y., T.S. Steenhuis, R.J. Glass, T.L. Richard, N.B. Pickering, W.J. Waltman, N. Bailey, M.S. Andreini, and J.A. Throop. 1988. The Flow of Pesticides Through Preferential Paths in Soils. New York's Food and Life Science Quarterly 18(1&2): 20-23. Cornell Univ., Ithica, NY

Richard, T.L., and T.S. Steenhuis. 1988. Tile Drain Sampling of Preferential Flow on a Field Scale. In: Rapid and Far-Reaching Hydrologic Processes in the Vadose Zone, P.E. Germann, Editor. J. Contaminant Hydrology 3:307-325,

Ritter, W. F., A.E.M. Chirnside, and R. W. Scarborough. 1987. Pesticide Leaching in a Coastal Plain Soil. ASAE Paper No. 87-2630. Amaerican Society of Agricultural Engineers, St. Joseph, MI.

-/103/32

A Computational Signal Processing Environment Using MATLAB®

Domingo Rodriguez, Engel Sánchez

Advanced Automated Image Analysis Group
Laboratory for Applied Remote Sensing and Image Processing
Electrical and Computer Engineering Department
University of Puerto Rico - Mayagüez Campus
Mayagüez, Puerto Rico 00681
http://exodo.upr.clu.edu/cecord

ABSTRACT
This paper presents an environment for computational signal processing applications using MATLAB®. The concept and main objectives of computational signal processing are introduced. The environment is then presented emphasizing its algorithmic approach and a methodology for analyzing, designing, modifying, and implementing multidimensional signal processing algorithms. An underlined mathematics al framework, based on Kronecker products, is discussed. Special attention is given to multidimensional fast Fourier transform based algorithms for radar signal processing applications. The basic features of the environment are discussed with some illustrations. Finally, some thoughts are given about the practical use of the environment as a tool aid in the mapping of algorithms to computational structures, auto-coding, and efficient rapid prototyping.

1.0 INTRODUCTION

This paper presents part of the work being conducted in the area of computational signal processing at the Electrical and Computer Engineering Department of the University of Puerto Rico. The work centers on the treatment of sensory data in order to obtain information important to a user. By sensory data we imply data collected from physical systems through the use of sensors. Great advances in sensor technology, communications and networking technology, and computer technology are demanding new theories, methods, and techniques to improve our understanding of our physical or sensory reality, For instance, better tools are needed for the tasks of manipulation, representation, visualization, and rendering when dealing with very large amounts of sensory data . One approach at improving on these tasks is to treat the data as signal sets carrying information which needs to be extracted. Treating the signals as elements in a set allows to study structures associated with the set and apply operator theoretic methods in an algebraic setting. This is the motivation for building a computational signal processing environment.

Computational signal processing uses operator methods, through algorithms, to assist at improving and enhancing the performance when effecting some of these tasks. Formally, computational signal processing deals with the algorithmic treatment of multidimensional] finite discrete signals in order to extract information important to a user. In this regard, a computational signal processing environment (CSPE) can be thought of as the aggregate of the following components: A set of input signals, a set of output signals, a set of operators, a set of composition rules for these operators, a set of actions rules for the operators to act on input data in order to produce output data, and a user interface.

In a CSPE we can implement operators for actions such signal correlation, convolution, domain transformation, decimation, interpolation, parameter detection and estimation, codification, and classification. The operators may be implemented using algorithms. The algorithms may take the form of software constructs, hardware constructs, or combinations thereof. In this paper we discuss the software implementation of operators through mathematical formulations and their mapping to computational structures in the form of algorithms. In this preliminary stage we are concentrating on operators for multidimensional correlation and convolution techniques, a large set of them implemented using multidimensional discrete Fourier transform operators. These correlation and convolution techniques are being used, in turn, in space-time-frequency analysis for synthetic aperture radar signal processing applications. We proceed to discuss in more details the concept of a computational signal processing environment and provide some examples of its usefulness.

2.0 COMPUTATIONAL SIGNAL PROCESSING ENVIRONMENT

As stated above, a CSPE is an aggregate of the following components: A set of input signals, a set of' output signals, a set of operators, a set of composition rules for these operators, a set of actions rules for the operators to act on input data in order to produce output data, and a user interface. Most CSPE operators are implemented through mathematical formulation of algorithms. To simplify the discussion, we will concentrate on multidimensional convolution algorithms; however, the concept can be extended to other types of operators as well. Convolution operators are important since they are used in applications such as discrete cross-ambiguity computations through two-dimensional operations of reflectivity densities and kernel imaging waveforms in radar signal processing [1],[2] (**see** Figure One). In this particular case, a CSPE would allow a user to access a set of input scenes or signals, apply a given convolution operator, and store or display the computed image or signal. The CSPE will also assist in the analysis, design, and implementation of the associated convolution algorithm in order to improve its computational performance. The assistance will be in the form of analyzing, designing, and implementing the functional primitives or expressions that conform a canonical mathematical formulation of an algorithm and its possible variants. Convolution operations can be performed indirectly through the use of the discrete Fourier transform (DFT)operator. We have developed a methodology, based on Kronecker or tensor products algebra, for the efficient implementation of discrete Fourier transform algorithms on computational structures[3],[4].

Figure Two below is a depiction of a CSPE. Algorithms associated with a particular operator are grouped into what we call computational signal processing frameworks. In the particular case of correlation, convolution, and other DFT-based operators. the language of Kronecker products algebra is used in a unified manner to identify similarities and differences among variants of a given algorithm. The optimization of software implementations on computational structures is sought through the study of the computational performance of each of the functional primitives that conform a given algorithm and the inherent parameters and attributes of the computational structure itself. At the present time, all algorithm implementations associated with the CSPE have been performed on individual workstations. Future implementations on heterogeneous distributed systems are envisioned.

3.0 CSPE RESULTS

The CSPE is based on the high-performance numeric computation and visualization software MATLAB®[5]. MATLAB® provided functions were used to a create graphical user interface and interact with a user. As stated above. the CSPE has been implemented on individual workstations. These include PC's and SUN Sparcstations. The CSPE was also tested for performance on SGI machines. Figure Three shows basic CSPE visualization tool-set for data representation and rendering. Figure Four provides and example of a CSPE operator set. In Figure Five we show basic CSPE data storage and retrieving capabilities. Choosing an operator from the menu wi 11 invoke this operator to act on the displayed data set, A sequence of operators can be applied on a data set in this manner in order to arrive. at a desired result. Hence, complex operators can be constructed via the composition of simpler ones. The CSPE retains a history file of the operators used in an interactive section. A user can request the MATLAB source code associated with a sequence of operators applied to the data set. This allows for a rough form of automatic code generation or autocoding. The autocode can be used by an interpreter or special purpose compiler to produce executable for a target machine or generalized computational structure. Having an autocoding procedure and a special purpose compiler associated with a CSPE wil I enhance efficient rapid prototyping efforts when dealing with algorithm implementati ons for scientific and engineering applications. A user, for inst ante, may want to improve a software algorithm, using the capabilities of the CSPE, for a given machine architecture, or he/she may want to obtain a stand alone hardware prototype.

4.0 CONCLUSION

We have presented a computational signal processing environment based on MATLAB® for data manipulation, representation, visualization, and rendering. The work seeks to contribute in the area of computational signal processing by providing results on the analysis, design, and implementation of multidimensional algorithms for radar signal processing applications through the use of' the CSPE. For the particular case of multidimensional DFT-based algorithms, we envision the CSPE being used to study algorithm computational performances on distributed heterogeneous computational structures. The language of Kronecker products algebra will play a crucial role in this process since certain inherent parameters of a structure can be incorporated into the algebra. It is this idea which allows us to establish a correspondence between a given mathematical formulation of an algorithm, in terms of functional primitives, and its associated computational constructs. As the constructs are improved, computing wise, the performance of that particular mathematical formulation also improves. We can fix a mathematical formulation an seek for a near optimal computational structure or fix a computational structure and search a near optimal mathematical formulation of the given algorithm for that particular computational structure. Also, an auto coding capability, associated with a special purpose compiler, is a step in the right direction on efficient rapid protot yping.

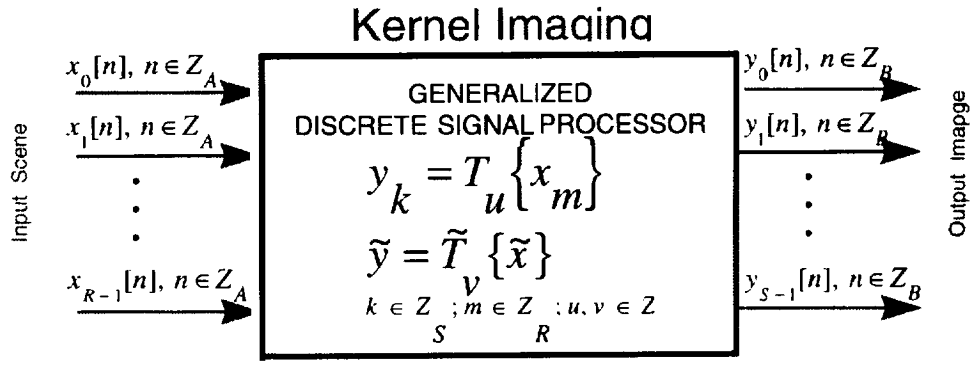


Figure One: CSPE implementation of kernel imaging convolution operation

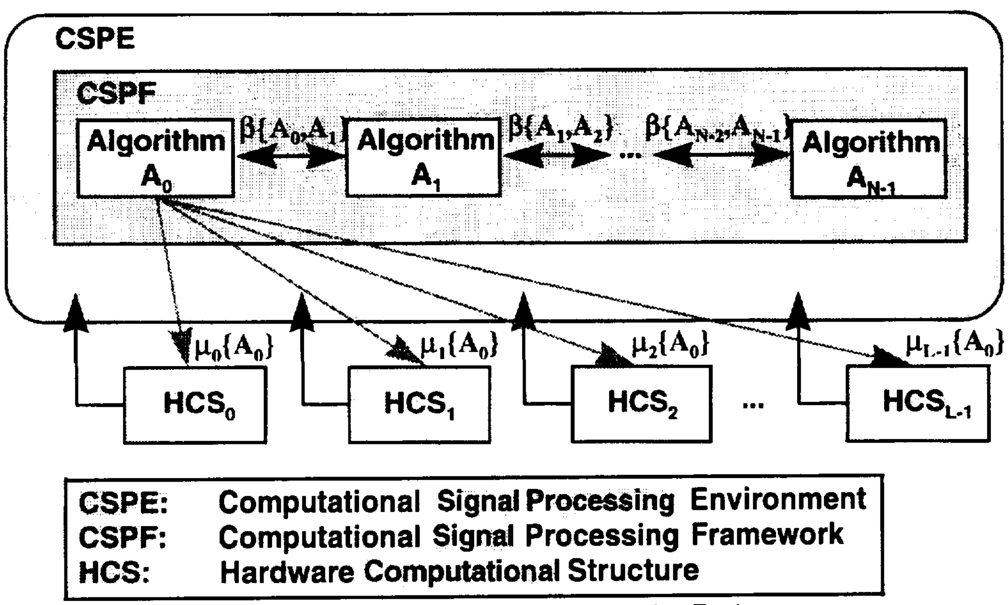


Figure Two: Computational Signal Processing Environment

REFERENCES

[1] R. E. Blahut, "Theory of Remote Surveillance Algorithms." Radar and Sonar, Part I, Springer-Verlag, New York, 1991.

[2]. D. Rodriguez, "Algorithms for Computing the Ambiguity Function on the IBM 3090 Supercomputer", 1991 International Conference of Society for Industrial and Applied Mathematics, Washington, DC, July, 1991

[3] J. Johnson, R. Johnson, D. Rodriguez, R. Tolimieri, "A Methodology for Designing, Modifying, and Implementing Fourier Transform Algorithms on Various Architectures. " Journal of Circuits, Systems and Signal Processing, Birkhäuser, Vol. 9, No. 4, 1990.

[4]. D. Rodriguez, "Tensor Product Algebra as a Tool for VLSI Implementation of the Discrete Fourier Transform," IEEE ICASSP '91, Toronto, Canada, May 1991,

[5]. MATLAB Reference Guide. The Math Works Inc., Natick, Mass., April 1995.

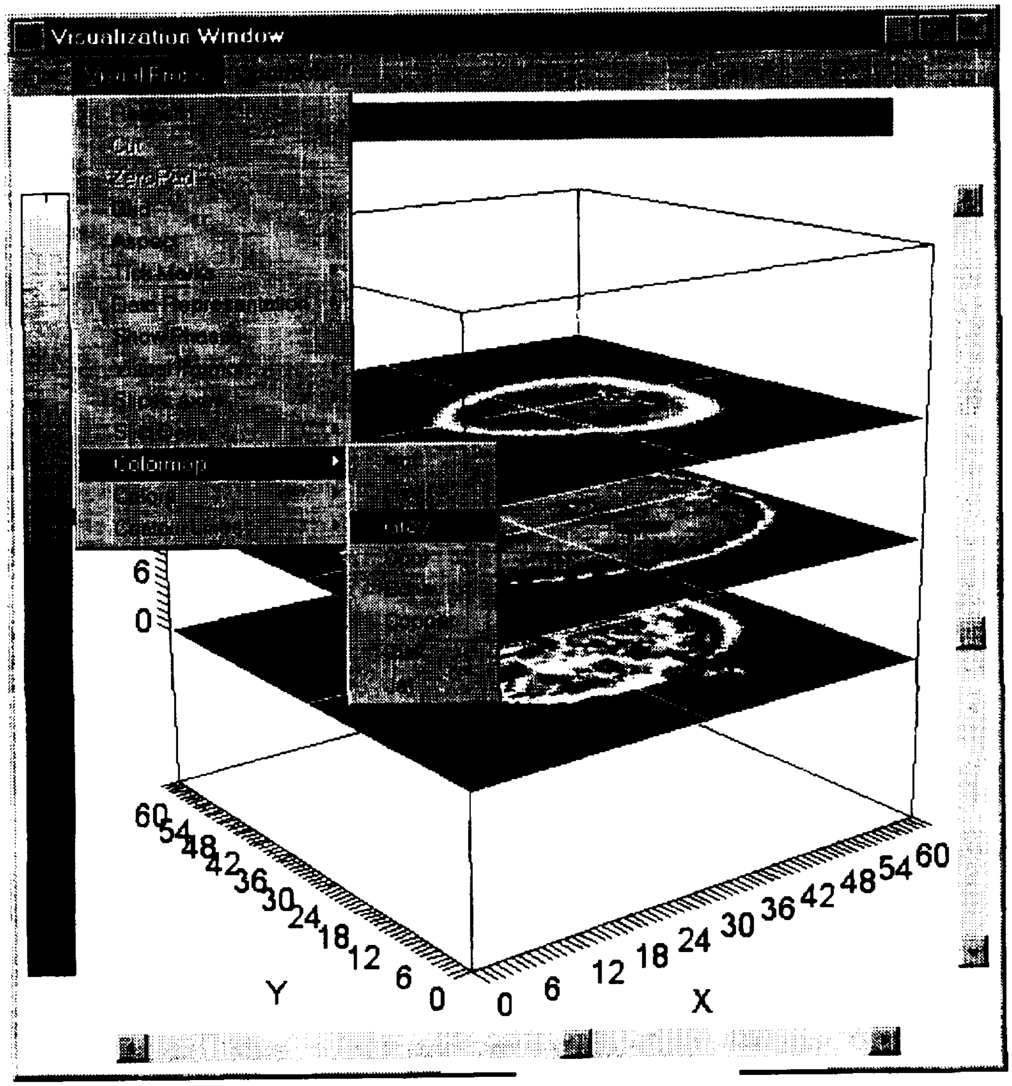


Figure Three: CSPE visualization tool-set.

Page intentionally left blank

//~~ NASA ~~ 5/07/43

Cluster Method Analysis of K.S.C. Image

Joe Rodriguez Jr. and M. Desai
Division of Engineering.

University of Texas at San Antonio
6900 Northloop 1604 West
San Antonio, TX 78249
joe@runner.utsa.edu

Abstract:

Information obtained from satellite-based systems has moved to the forefront as a method in the identification of many land cover types. Identification of different land features through remote sensing is an effective tool for regional and global assessment of geometric characteristics. Classification data acquired from remote sensing images have a wide variety of applications. In particular, analysis of remote sensing images have special applications in the classification of various types of vegetation. Results obtained from classification studies of a particular area or region serve towards a greater understanding of what parameters (ecological, temporal, etc.) affect the region being analyzed. In this paper, we make a distinction between both types of classification approaches although, focus is given to the unsupervised classification method using 1987 Thematic Mapped (TM) images of Kennedy Space Center.

I. Introduction:

The primary objective of image classification is to identify, as a unique gray level, the features occurring in an images in terms of the type of land cover these features actually represent. Picture elements, pixels, within an image represent the smallest unit of spatial area on the ground for which data is collected. Image analysis is done to provide a quantitative analysis of pixels for which, using computer based algorithms, they are counted for area estimates and identified based on their numerical properties.

Basic to the understanding of multispectral classification is the concept of the spectral signature or spectral response of an object on the ground. The spectral response for a given object is a measure of the amount of electromagnetic radiation it reflects as a function of wavelength. This quantitative measure of the reflected electromagnetic radiation sampled in a series of different wavelength bands produces a unique response called a signature. Therefore, the objective of classification becomes recognition of unique pixel signatures [2]. The governing idea is to automatically categorize all signatures in an image into special land cover classes, more commonly know as themes.

Remote sensing data sets exits as a or way to integrate spatially heterogeneous responses into a more easily measurable format by quantifying them at a specific scale (e.g. 10, 20, or 30m). For instance, Landsat Thematic Mapper (TM) sensors have a spatial resolution or pixel size of 30m, which represents a 30m x 30m area on the ground. In an area of heterogeneous land covers, spectral responses for different objects within a pixel will be averaged or aggregated into a composite spectral response for any particular pixel that falls over a specific area on the ground.

Thus the multispectural domain exists as both an additional axis of information available for analysis and as any integrating factor of scale-related phenomena [1,3].

H. Classification Analysis:

Multispectral classification is an information-extraction process that analyzes the spectral signatures determined in a region and then assigns pixels into categories based on similar signatures obtained in the entire image. There are generally two types of classification approaches used, supervised and unsupervised.

Supervised classification procedures are part of the essential tools used to extract quantitative information from remotely sensed data. In this type of approach, the analyst defines on the image a small area, called a training site, which is representative of each terrain category, or class. Then spectral values for each pixel in a training site are used to define the decision space for that class. After each training site is defined, the computer algorithm then classifies all the remaining pixels in the scene accordingly.

The second classification approach, for which we focus our analysis, is called unsupervised. Unsupervised classification is a method which examines a large number of pixels and divides them into a number of classes based on natural groupings present in the image. Unsupervised Classification is performed most often by using clustering methods to assign each pixel in an image to spectral classes, of which a user has no previous knowledge. Unlike supervised classification, unsupervised classification does not require analyst-specified training data (previously acquired data of the scene being analyzed). This procedure can be used to determine the number and location of the spectral classes into which the pixels are assigned. Finally, using the existing information from site photos, visits, and maps the resulting classification can then be identified [2].

The unsupervised approach to image classification always requires the classifier, the algorithm used to carry forth the pixel analysis, to learn or cluster. Clustering techniques are useful for image segmentation and for classification of raw data, for which there is no previous knowledge, to establish classes. Statistical techniques can be utilized to automatically group an n-dimensional set of observations in their natural spectral classes. Therefore we use clustering techniques to define a set of feature points, pixels, in the region being analyzed for which their is a large density compared to the density of features points in the surrounding region.

For our analysis we adopted a computer based clustering method, ISOCLUS [8], which is an iterative statistical method for clustering of feature points. This clustering method is based on the Isodata algorithm. In the analysis, we first assumed the number of clusters, K. Next, the partitioning of the data is done such that the average spread or variance of the partition is minimized. Let $\mu_k(n)$ denote the kth cluster center and the nth iteration and $\mathbf{R}_k$ denote the region of the kth cluster at a given iteration. Initially, we assigned arbitrary values of $\mu_k(0)$. At the nth iteration we took one of the data point $\mathbf{x}_i$ and assigned it to the cluster whose center is closest to it, that is,

$$\mathbf{x}_i \in R_k \iff d(\mathbf{x}_i, \mu_k(n)) = \min [d(\mathbf{x}_i, \mu_k(n)]$$

where $d(\mathbf{x},\mathbf{y})$ is the distance measure used. Then we recompute the cluster centers by finding the point that minimizes the distance for elements with each cluster. Thus,

$$\mu_k(n+1): \sum d(\mathbf{x}_i, \mu_k(n+1)) = \min \sum d(\mathbf{x}_i, \mathbf{y}), \qquad k = 1, \ldots, K.$$

The procedure is repeated for each $\mathbf{x}_i$, one at a time, until the clusters and their centers remain unchanged. If $d(\mathbf{x},\mathbf{y})$ *is* the Euclidean distance, then a cluster center is simply the mean location of its elements. If K is not known, we start with large values of K and then merge to K-1, K-2, . . . clusters by a suitable cluster-distance measure. [3,4].

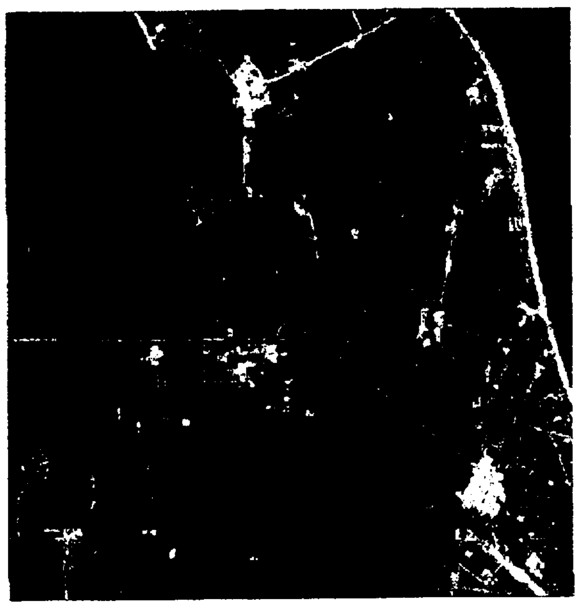


Figure 1, Original 1987 TM Image of KSC

An original portion of a 1987 TM image (using BANDS 1, 2, 3) taken of Kennedy Space Center for which we used in our analysis is represented in ***Figure 1.*** The objective is to identify major land cover classes, primarily vegetation, using the ISOCLUS algorithm. We suspect that using this type of classification scheme will yield at least three different distinctions: manmade structures, water, and vegetation.

III Classification Results:

Figure 2. shows the image after the classification analysis has been done using the ISOCLUS algorithm. ***Figure 3*** represents color composite of ***Figure 2,*** this was done to highlight the differences in contrast in order to more easily analyze and review the results. Careful examination of the results shows that the ISOCLUS algorithm successfully identified the those regions corresponding to water, manmade structures, and vegetation. The signature given for water is clearly given by the major black regions in ***Figure 2 &3.*** Manmade structures are easily depicted from the ***Figure 2*** and are given a signature in the white to light-level gray areas. The mid-level to somewhat dark shades of gray in ***Figure 2*** and blue-green regions in ***Figure 3*** correspond to vegetation. It is important to point out that classification of vegetation occurred in two different types of regions, meaning that the ISODATA algorithm recognized healthy mainland vegetation

as well as vegetation growing in lower wetland areas. Close observation of the upper left hand portion of *Figure 2* illustrate this

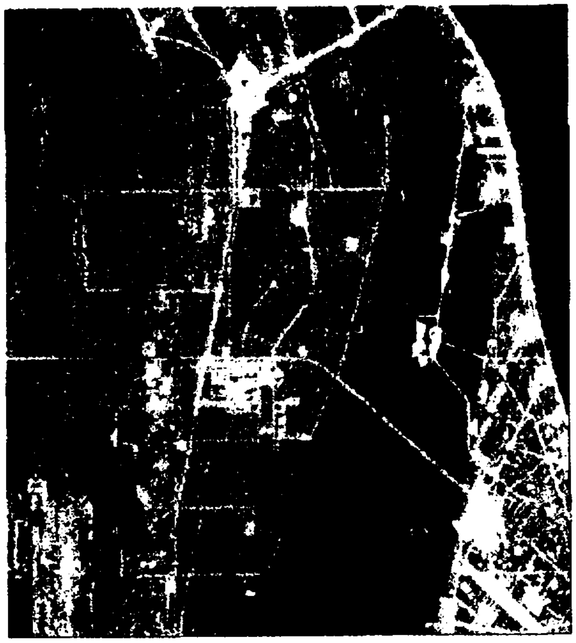


Figure 2. Image after Clustering Analysis

Figure 3. Color Composite of Clustered Image

IV. Conclusion:

For the analysis of these images using the unsupervised technique vegetation signatures identified were only put into general class and were not discretely defined exclusively by vegetation type. Results obtained form these K.S.C. image primarily serve as a stepping stone for more extensive analysis using more complicated techniques. We will use the these results and incorporate them with fuzzy logic analysis to obtain an exclusive distinctions between vegetation types.

V. Acknowledgments:

This work was partially supported by a grant form NASA No. NAG 10-0155 and by a grant form the University of Texas at El Paso No. NCCW-0089

VI. References:

[1] D. Quattrochi and R. Pelletier *Remote Sensing for analysis of landscapes : An Introduction*, 1991.

[2] F.F. Sabin. Jr., *Remote Sensing Principles and Interpretation*. New York: W H. Feeman & Co. 1987.

[3] A.K. Jain. *Fundamentals of Digital Image Processing*. New York: Prentice Hall. Inc. 1989.

[4] T.M. Lillesan and R.W Kiefer. *Remote Sensing and Image Interpretation*. New York: Wiley 1979.

[5] B. Wood. L. Beck. S. Dister. B. Lobitz. and V. Ambrosia. *"Readings in Remote Sensing. Geographic Information Systems and Related Technologies."* NASA Ames Research Center. 1996.

[6] A. H. S. Sloberg. A. K. Jain. and T. Taxt. *"Multisource classification of Remotely Sensed Data: Fusion of Landsat TM and SAR Images"*. IEEE Trans. Geosci., Remote Sensing. vol. 32 no. 4. p 768-777

[7] L. E. Pierce. F.T. Ulaby. K. Sarabani. and M.C. Dobson *"Knowledge-Based Classification of Polarimetric SAR Images"*. IEEE Trans. Geosci., Remote Sensing. vol. 32 no. 5. p 081-1086.

[8] *Using PCI Software Vol. II*. PCI Inc. Version 5.3. Copyright 1994.

URC97110

A HYDROGEOLOGIC AND REMOTE SENSING INVESTIGATION OF AQUIFER CONTAMINATION BY NITRATES FROM *AGUAS NEGRAS IN* CHIHUAHUA, MEXICO

Rodríguez-Pineda J. A., Pingitore N. E., Perez A., Penn B. S., Keller G. R.
All at the Geological Sciences Dept. at the University of Texas at El Paso
El Paso TX, 79968

I. Introduction

Chihuahua, Chih., Mex., with a population of around 700,000 inhabitants, has released untreated sewage for decades to the intermittent Chuvíscar River, currently at a rate of 3,100 ls^{-1}. Downstream the river drains the Tabalaopa-Aldama Valley conducting mainly wastewater. This raw sewage is used by local farmers as a source of water and fertilizer for approximately 50°/0 of the agricultural areas in the valley. The other 50 °/o are irrigated with groundwater.

This aquifer supplies 600 ls^{-1} of groundwater, through 15 deep wells, for almost 300,000 people in Chihuahua, Aldama, and several rural communities. In some areas nitrate pollution is above the Mexican regulation (22 mg l^{-1} as NO_3^- or 5 mg 1-* as N.)

We are analyzing satellite imagery to detect spectral response differences between agricultural crops irrigated with *aguas negras* and the same type of crops irrigated with groundwater. If a spectral difference does exist between them, it could be correlated with aquifer pollution degree. This difference could be used to detect agricultural areas overfertilized with natural or comercial fertilizers that are threatening the groundwater reservoirs below them.

II. Research area background

1. Location

The Tabalaopa Valley is located between the cities of Chihuahua and Aldama in the central part of Chihuahua state, in northern Mexico (Fig.1). The geographical coordinates are latitude 28°42'- 29°42' N and longitude 105°56'-106°08' W. The valley is 25 km long and 13 km wide, and covers an area of approximately 325 km^2.

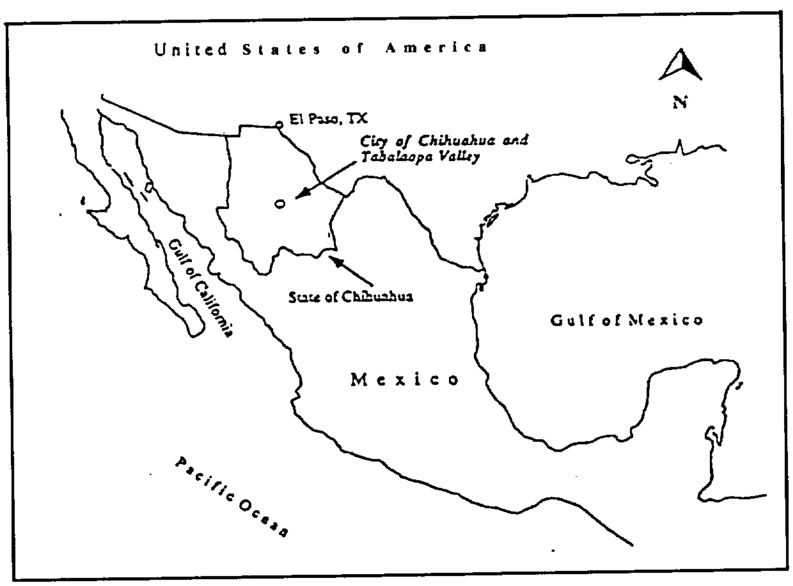


Fig. 1. Location of the study area.

2. Climate

With a semi-arid climate, due to its location at the west boundary of the Chihuahua Desert, the area experiences a wide range of temperatures normally between -5 to 40° C through the year,

with a mean of 18.20 C. Summer is the rainy season and annual rainfall averages 395 mm y-l. The potential evaporation value varies from 62 mm/month during winter to 253 mm/month during summer, with mean of 148 mm/month (Comision National del Agua, 1993). The annual mean water deficit in the zone ranges from 500 to 600 mm y-l (Carta de Evapotranspiracion y Deficit de Agua, S.P.P.).

3. Geology

Geologically the Tabalaopa Valley is part of the Basin and Range Province. This province is a broad zone of continental rifting characterized by extensional fault-block mountains and deep sediment-filled basins. The valleys were formed by regional extension during the late Tertiary by sub-parallel normal faults. Later, these grabens were filled with alluvial sediments that function as excellent water reservoirs. The Tabalaopa Valley, like the Hueco Bolson in the area of El Paso-Juarez, is a typical example of such a reservoir in the Basin and Range Province .

The regional stratigraphic section consists of Precambrian rocks (metagranites, amphibolites, and minor gneisses) that crop out in the Rancho Los Filtros area 15 km NNW of Aldama (Mauger and McDowell, 1983). The rest of the stratigraphic section consists of Cretaceus carbonates and evaporates unconformably overlain by Tertiary volcanic rocks and volcaniclastic sediments (Maldonado and Megaw, 1983). This stratigraphic sequence is observed in the ranges surrounding the Tabalaopa Valley, Sierra Sacramento on the west boundary, and El Cuervo-Peña Blanca to the north. To the east of the Tabalaopa Valley is the Santa Eulalia range, and finally, to the south is the Pastorias Range.

The Tabalaopa basin is composed of Quaternary, Pleistocene and Holocene alluvial non-consolidated sediments. The aquifer occurs in a sedimentary environment dominated by alluvial fans (Chavez, 1993). These sediments are composed mainly of gravel, boulders, sand, and some lenses of clay observed during well drilling.

4. Surface hydrology

The city of Chihuahua releases approximately 3,1001 S-l of untreated sewage to the dry bed of the intermittent Chuvíscar River. Due to the semi-arid conditions of the area and the Chihuahua and El Rejon dams, this river does not have a water flow base. Without the city sewage discharge the Chuviscar River would remain dry during most of the year. Only during the rainy season does a mixture of *aguas negras* and rainwater flow. The city sewage is composed of 90°/0 domestic water, 6% from industry and 4% from commercial activity (Calderon, 1994).

This river, with a slope mean of 0.4%, drains the 23 km long of Tabalaopa Valley. Along the river's path almost 60,000,000 m^3 y-l of waste water (Chávez et al., 1993) is diverted by local farmers for agricultural irrigation. This value corresponds to 61% of the total sewage released by the city and entering the valley. The conduction of sewage toward the agricultural areas is through a series of unlined channels.

A water balance in the valley shows that the only significant water loss is due to evapotranspiration. Waste water that comes into the valley via the river infiltrates or evapotranspires, but none spills into the next valley, except during extreme precipitation events.

The Chuvíscar River is considered to be the primary recharge source for the aquifer (Martinez et al., 1978; Osuna, 1991; Chavez et al., 1993). Nevertheless, the almost 2,000 hectares of agricultural lands irrigated with untreated sewage and the unlined irrigation channel system use 610/0 of the raw sewage that the river conducts. This number suggests that irrigation return can be as important as the river as a source of aquifer recharge.

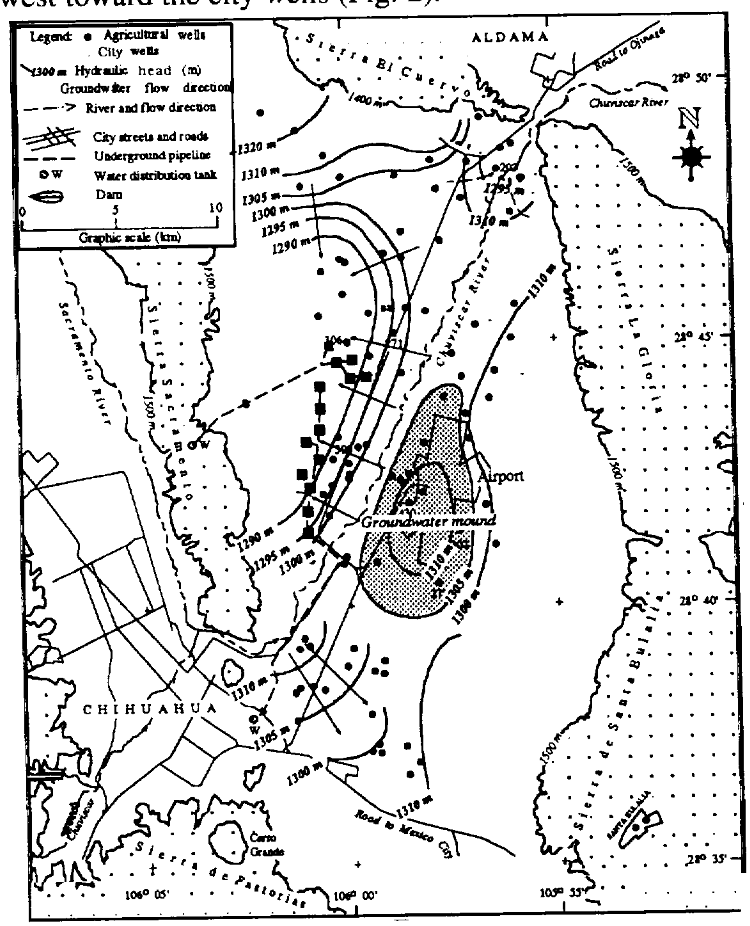
The rest of the surface hydrology consists of small creeks associated with hill drainage around the valley. These creeks conduct water only during the rainy season, apparently contributing to just

a small part of the annual water balance in the basin.

5. Subsurface hydrology

The Tabalaopa-Aldama aquifer is in a porous sedimentary environment dominated by alluvial fans and fluvial deposits with some intercalations of clays (Chavez et al, 1993) with transmissivity values range from 10 to 2000 m d-l. Two wells with specific yield show values of 0.16.

Due to surface topography, the groundwater table depth ranges from -120 to -5 m from S to N. Before the aquifer exploitation, the groundwater flow was from S to N; this trend has been modified by the 15 deep city wells located in the west zone (Fig. 2). Now, almost the entire aquifer responds to the cone of depression generated by the city wells, and the groundwater flow is to the west toward the city wells (Fig. 2).



The Tabalaopa aquifer supplies 25% (22 million m^3 y-l) of potable water for the city of Chihuahua. Additionally, the region has nearly 90 agricultural wells, 8 wells for water supply of rural communities, and several shallow wells along and close to the river edges.

A groundwater mound is formed beneath the agricultural zones irrigated with sewage (Fig. 2). This dome is created mainly by the excessive amount of sewage used for irrigation, and, possibly by river infiltration.

Fig. 2. Hydraulic heads (meters above sea level) and direction of groundwater j-low

6. Agricultural activity

Around 2,100 hectares are irrigated with 60 million m^3 (61%) of raw sewage per year, and the rest is irrigated with groundwater and no application of fertilizers. Irrigation by flooding is the technique used in Tabalaopa Valley, mainly at those areas irrigated with raw sewage. This practice enhances the seepage of untreated sewage into the soil (Rodriguez, 1994).

The main crops irrigated with sewage are alfalfa, sorghum, and rye grass. The areas irrigated with raw sewage do not require the use of fertilizers because the raw sewage supplies both moisture and nutrients to the crops.

III. Research methodology

1. Groundwater sampling

A three-year groundwater sampling program was established to detect nitrate and contaminant trends due to river infiltration and irrigation return. This program was undertaken during the summer season when most of the agricultural wells were pumping. Some samples were obtained from shallow wells, and from the river itself

1.1 Nitrate ion

Nitrogen is an important element in waste waters that ranges from 20 to 85 mg l^{-1} in a typical raw municipal sewage (Feigin et al., 1991). High concentrations in groundwater suggest contamination by sewage or fertilizers. It can appear as nitrate (NO_3^-), nitrite (NO*-), nitrous oxide (N_2O), ammonium (NH_3^-), and as nitrogen gas (N_2). The nitrate ion, which is considered an environmental hazard, is the last oxidation stage of nitrogen exposed to aerobic conditions . Due to its negative charge this ion is not retarded by the soil; consequently, nitrate travels at almost the same speed as the groundwater. These characteristics help in the use of nitrate as a tracer to define polluted zones.

Nitrate ion was analyzed immediately at the well location after the sample was taken using the technique of cadmium reduction with a Hach surface water kit. This avoids further oxidation of nitrogen compounds and consequent increase in the nitrate concentrations. Later, a comparative analysis between Hach field technique and the ion chromatography technique, which is the standard method for nitrate determination, was done in the laboratory to validate nitrate results.

2. Spectral analysis of agricultural crops

Data from the Landsat satellites have been available since the 1970's. These have been used extensively in studies of dryland vegetation and land cover (Millington et al., 1994). In this study two Landsat TM images, 180 km by 185 km and 28.5 m nominal resolution, taken in March 1986 and June 1992, are used to identify the spectral attributes of crop plots irrigated with untreated sewage and those irrigated with groundwater.

The Tabalaopa Valley was grouped in two areas. The south part is irrigated with sewage and the north area is irrigated with groundwater. The spectral response generated by reflected light from these two areas could be different. Nitrogen is a macronutrient for plants, and sewage is an important source of N. Consequently, the crops in the Tabalaopa Valley irrigated with sewage have excess nitrogen and could show different reflectance characteristics then those areas irrigated with groundwater. This situation maybe similar to overfertilized agricultural areas with industrial nitrogen fertilizers, such as urea (46.6% nitrogen), that represent a risk for aquifers.

2.1 *Secular changes in extent of agriculture*

The agricultural development of the Tabalaopa valley with time is not well documented. This condition may be proportional to aquifer degradation. Nitrate information from 1978 (Martinez et al. 1978) and our database will assist in developing the relation between valley agricultural development and aquifer contamination.

Analyzing aerial photographs from 1976, 1994, and TM satellite images from 1986 and 1992, we were able to determine the approximate agricultural area extent irrigated with sewage for each of these years. The north area of the valley, irrigated with ground water, was not considered in any of the area calculations.

The southern part of the Tabalaopa Valley is irrigated with *aguas negras*. Using the aerial photographs from 1976 and 1984 the area was subdivided into specific geometric sections. The area

of each of each of these subdivisions was calculated and added to obtain a total approximate area of irrigation with sewage.

Two Landsat TM satellite images, taken in March of 1986 and June of 1992, were used to obtain the area cultivated using sewage. To obtain this area we used two different methods. The first one was the application of the same technique applied to the aerial photographs. The second method involved the selected regions of interest (ROI). From each ROI the number of pixels was obtained. Knowing pixel dimensions of 28.5 m by 28.5 m the total area was calculated based on the number of pixels in each ROI.

IV. Preliminary results

1. Nitrate results

Th itrate concentrations (Fig. 3) clearly show that the aquifer is polluted by the seepage of irrigation return and river infiltration. Some areas of the Tabalaopa aquifer exceed the nitrate concentration of 22 mg l^{-1} as NO_3^- established by Mexican regulation (Articulo 213. Diario Oficial de la Federacion Jan. 18, 1988). Nitrate concentration contours show three main highs at the middle of the valley (Fig. 3). One is in the south part of the valley. The second is below the main area irrigated with sewage. The third is along the river path from south to north. The general trend of these polluted zones is in the direction of the city wells due to the large cone of depression created by them.

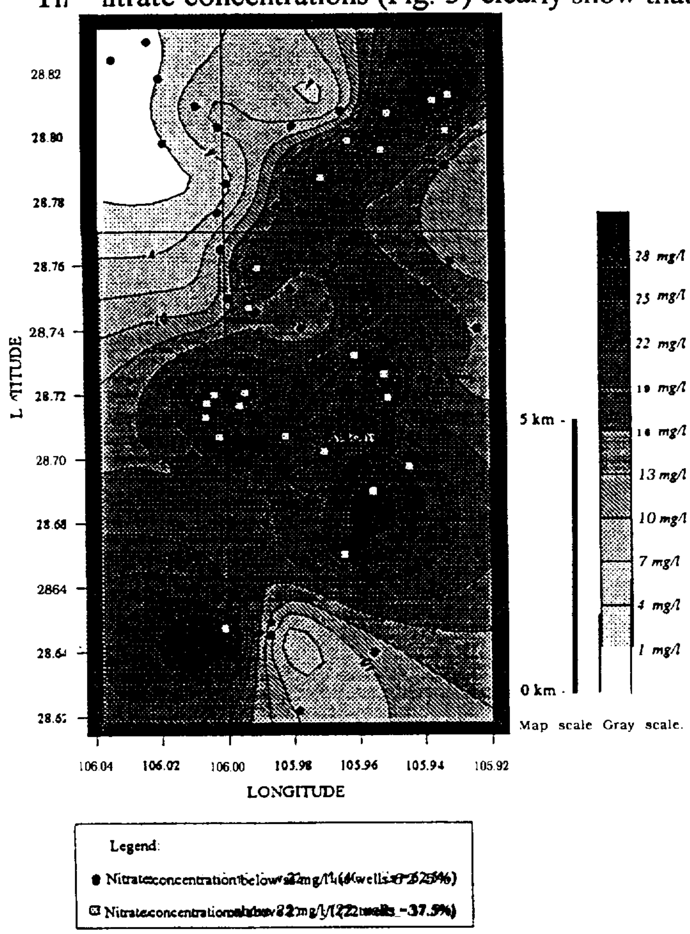


Fig. 3. Spatial distribution of nitrates in the Tabalaopa aquifer (1995 data).

2. Agricultural Secular changes

The areas irrigated with sewage have increased with time. From the aerial picture taken in 1976, We extimate that 1,000 hectares were irrigated with sewage water. In 1986 the satellite imagery showed an increase to about 1,640 hectares. Finally, the 1994 aerial photograph displayed 2,180 ha.

The time period analyzed covered eighteen years, and the total increment of agricultural areas between 1976 and 1994 was 1,180 ha., an increment average of 65.5 ha per year. The different images show that most of the expansion of sewage irrigated agricultural land ocurred in the first decade. Addition of agricultural areas irrigated with sewage occurred mainly south part of the airport. Another important area was developed between the Chuviscar River and the airport.

V. BIBLIOGRAPHY

Blount J. G., The geology of the Rancho los Filtros area, Chihuahua, Mexico., Guidebook for the 1983 field conference, El Paso Geological Society, October 1983. p. 157.

Calderón M. Evaluation del impacto ambiental de las descargas de aguas residuals en la ciudad de Chihuahua., 1994. Master's thesis. Fat. de Ingeniería, Universidad Autónoma de Chihuahua.

Chávez A., Chávez R., Blanco R. Modelación matemática del transport de contaminantes en el acuífero de Tabalaopa-Aldama, Chihuahua. 1993. Fat. de Ingeniería, Universidad Autónoma de Chihuahua.

Chávez R. Hidrogeología física y química de la portion centro-occidental del vane de Tabalaopa-Aldama, Chihuahua. 1993. Master's thesis. Fat. de Ingeniería, Universidad Autónoma de Chihuahua.

Feigin A., Ravina I., Shalheve J. Irrigation with treated sewage effluent: managment for environmental protection. Springer-Verlag Berlin Heidelberg 1991.

INEGI., Carta de evapotranspiración y deficit de agua, carta Chihuahua, scale 1:1,000,000. Edited by the Secretarial de Programación y Presupuesto, 1983.

Maldonado D., Megaw P. Geology of the Santa Eulalia Mining District, Chihuahua, Mexico., Geology and mineral resources of north-central Chihuahua. Guidebook for the 1983 field conference, El Paso Geological Society, October 1983. p. 367.

Martinez P.,De la Rosa, A., Monsivais, H., Charles, J., Ramírez, A., Bernal, M., Morales, A. Estudio Geohidrológico del Vane de Chihuahua, primers etapa. Serie de Investigaciones Geohidrológicas No. 100, Fat. de Ingeniería, Universidad Autónoma de Chihuahua, 1979.

Mauger R. L., The geology and volcanic stratigraphy of the sierra Sacramento block near Chihuahua city, Chihuahua, Mexico. Geology and mineral resources of north- central Chihuahua. Guidebook for the 1983 field conference, El Paso Geological Society, October 1983. p. 137.

Mauger R. L., and McDowell F. W., Grenville-Age Precambrian rocks of the los filtros area near Aldama, Chihuahua, México.

Millington A. C., Wellens J., Settle J. J., and Saull R. J. Explaining and monitoring land cover dynamics in drylands using multi-temporal analysis of NOAA AVHRR imagery, in Foody G., Curran P., Environmental remote sensing from regional to global scales, John Wiley and Sons, 1994.

Pingitore N. E., Geological, hydrological, environmental, geochemical, and agricultural investigations in the Cd. Chihuahua region, Chihuahua, México. Internal

5111/-7

THE USE OF DECENTRALIZED CONTROL IN THE DESIGN OF A LARGE SEGMENTED SPACE REFLECTOR

by

Helen Ryaciotaki-Boussalis
Maj Mirmirani
Khosrow Rad
Mauricio Morales
Efrain Velazquez

California State University, Los Angeles
5151 State University Dr.
Los Angeles, CA 90032

Anastasios Chassiakos
Jose-A lberto Luzardo

California State University, Long Beach
1250 Bellflower Blvd.
Long Beach, CA 90840-5602

ABSTRACT

The 3-dimensional model for a segmented reflector telescope is developed using finite element techniques. The structure is decomposed into six subsystems. System control design using neural networks is performed. Performance evaluation is demonstrated via simulation using PRO-MATLAB and SIMULINK.

1. INTRODUCTION

Future astronomical space missions will require high-performance optical systems. Such optical systems will necessarily have large apertures for high-precision performance making the size of the reflectors very large. Due to practical considerations such as time, cost and complexity of fabricating large-aperture telescopes, as well as launch vehicle size. weight and power constraints, future telescopes wi ll be made of precision-segmented reflectors. A segmented mirror reflector consists of mirror panels which, when formed together, become a parabolic primary m irror that magnify the images from space. These mirror panels are easily manufactured, and deployable in space when the proper orbit is achieved.

With the advantages of the segmented reflector telescopes, shortcomings are also present. The large size of these telescopes make the structure flexible to external forces such as thermal fluctuations and solar disturbances. It becomes apparent that development of control concepts and extensive performance evaluation via simulation in an environment characterized by various dynamic disturbances is a necessity. The control design challenge is to make the segmented reflector perform as a monolithic reflector. This is done via high-precision figure control which would maintain the surface of the reflector to within a specified tolerance of the calibrated reference position.

Due to the large size of the telescope structure, it becomes apparent that control system design based on conventional methods is exceedingly difficult. The mathematical model of the structure involves hundreds of states and use of centralized control ler could not accomplish figure control according to specifications. Decentralized control has been used by several researchers in the recent years to overcome the difficulties due to the dimensionality problems that arise when dealing with large-scale systems [5]. This approach has been taken here to accomplish control of the segmented reflector. Specifically, the structure has been decomposed into six smaller-order subsystems. The decentralization is performed physically by slicing the structure vertically and isolating every mirror panel along with its associated structural member.

The control design has been performed using PID [1], pole-placement, H_2 (LOG), H-infinity and neural networks control. The neural network based control is discussed in the present paper. The control law is developed at the local level and performance evaluations are performed utilizing the control law of the isolated subsystems and the interaction among them.

Section 2 contains the mechanical and structural design of the structure, section 3 lists the control system requirements, section 4 contains the mathematical model, section 5 describes the decentralization of the structure, section 6 contains the neural network based control design, section 7 contains the computer architecture of the structure, and section 8 contains the summary of the results.

2. MECHANICAL AND STRUCTURAL DESIGN

The Control and Structures Research Laboratory (CSRL) segmented reflector testbed is an experimental apparatus capable of addressing the technical challenges presented by a complex three-dimensional structure such as a large segmented optical system. To validate both control strategy and implementation for realistic systems, a set of requirements have been used to design the testbed [2]. These requirements are based on various missions within NASA's Precision Segmented Reflector Program (PSR) [1], They are also based on the requirements used for similar projects, including JPL's PSR and Lockheed's ASCIE testbeds. Since the CSRL testbed is a control-system oriented instrument, requirements are selected mainly to demonstrate a high level of disturbance attenuation rather than optical performance.

PERFORMANCE: Performance of the testbed is required to be of comparable quality to that of an actual system. The testbed is designed to perform the essential functions needed for the various system missions including static and dynamic segment alignment, fine pointing and vibration suppression.

DYNAMICS: The structure is designed to approximate the fundamental dynamic characteristics of a three-dimensional large structure, i.e. low-frequency modes, high modal density, and global mode shapes that properly reflect the coupling of the sub-elements of the structure. The system is designed to accommodate interdisciplinary experiments in validation of control algorithms, CSI, optics, electronics, actuators, sensors, and distributed multiprocessor design and implementation. The system is designed to demonstrate physical and mathematical decentralization and accommodate development of control algorithms related to decentralized control technique.

2.1 CSRL system description. Figure 1 is a schematic illustration of major features of the testbed, including the primary and secondary mirrors, the actuators, the edge sensors. The active optical elements are the primary-mirror segments which interact dynamically with the actuators, sensors and the supporting structure in an integrated way. The primary mirror is a 2.63 m diameter dish supported on a lightweight, flexible truss structure. The optical system emulates that of an f / 2.4 m Cassegrian telescope. The major components of the CSRL testbed are discussed bellow.

STRUCTURE: One of the most fundamental design goals has been a strong, light-weight truss structure whose structural-dynamic characteristics are representative of a large, flexible space-borne system. These include low frequency modes, high modal density and global mode shapes that properly reflect the coupling of the sub-elements of the structure [2]. Therefore, the requirements for the design of the truss, in addition to the dimensional constrains, included a careful trade-off between the need for the structure to support itself in the 1-g laboratory environment versus the need to keep the frequency of the first mode as low as possible. Multi-criteria optimization technique based on Pareto optimality concept was employed to accomplish this objective. The overall dimensions are 2.275 m across and 0.580 m thick. The structure is made of nine groups of stainless steel truss elements ranging in size from 0.921 m to 0.414 m. There are 60 elements and 21 nodes. The truss is supported on a specially designed isolation platform.

SEGMENTED PRIMARY MIRROR : The CSRL primary mirror is designed to emulate the critical properties of a real segmented mirror. These properties include segmentation geometry, inter-segment spacing, segment mass, inertia and stiffness, and optical focal ratio. The seven segment primary mirror consists of a ring of six actively controlled hexagonal segments surrounding a fixed center segment that acts as a reference. Because the testbed is control-system oriented, and because of difficulty and added expense of fabrication of actual optical-quality segments made from glass, the segments will be fabricated from flat honeycomb aluminum plates. The active segments are attached to segment-positioning actuators with special three-degree-of-freedom flexures. These flexures permit individual segments to have two rotational degree of freedom (tilt) and one transitional degree of freedom (piston). Each segment is controlled by three actuators and the entire primary has a total of 18 actuators. The relative displacement

between the edges of adjacent segment unmeasured by an ensemble of 24 edge sensors. The edge sensors provide information about a segment's relative displacement as well as absolute displacements from the fixed center reference segment.

SEGMENT-POSITIONING ACTUATORS: Use of high performance segment-positioning actuators are the key to precision control of the CSRL testbed primary mirror. These actuators must have extremely low noise level, be able to generate substantial force over a wide mechanical range and support the weight of a segment in a 1-g field. They must also have a bandwidth sufficient to accommodate the spectrum of expected disturbance and to support robust control of the system. In addition, they should be able to actuate free of friction, and minimize thermal energy dissipation. These actuators must be fitted with collocated positioning sensors and/or accelerometers, be modular and compact in size and easily interface with the structure. Because conventional actuators are unable to meet one or more of the above performance requirements a voicecoil design has been especially developed for the CSRL testbed. These actuators are being fabricated by Northern Magnetic, Inc. in Southern California. Some of the mechanical features include especially designed disk flexures instead of conventional bearings and an off load spring is to minimize actuator force requirements and thermal energy dissipation when the actuators are holding the weight of the segment. The actuators have a bandwidth of 0-150 Hz., positioning resolution of 0.1 micrometer and a maximum force output of 54 Newtons.

ACTIVE SECONDARY MIRROR: The CSRL testbed secondary reflector design consists of a 12.5 in mirror supported by a tripod that is attached to the primary truss at three points. The mirror is designed to provide two-axis, active beam-steering control. An active closed-loop control system is being designed that is capable of aligning the secondary to the focal plane, removing all relative angular motion between the secondary and the reference center segment of the primary structure. The secondary mirror is supported by isolation springs attached to the secondary structure. Three reluctance actuators located 120 degrees from each other provide for three degrees of freedom (tip, tilt, piston) motion. Three position sensors are used in combination with the actuators to control the position of the mirror.

2.2 **Structural optimization.** Pareto optimality concept [4] was used to design a structure which represents the "best" trade off between flexibility and strength. The panel surface distortion due to gravity as measured by the RMS distortion of the upper truss nodes and total mass of the structure were selected as two criteria for optimization. A set of Matlab subroutines which use finite element data to arrive at an optimal solution were developed. Two sets of geometric parameters were allowed to vary at specified increments to obtain variations of the baseline structure. RMS values and total mass were calculated for each resulting structure. The structures with different geometric characteristics plotted as a collection of points in the objective space show a Pareto optimal pattern on the boundary of the region (Figure 2). Only points on the left of the dotted area from point A to point C represent Pareto optimal solutions. Point A represents the structure with minimum RMS surface distortion, while point C represents the minimum weight structure. The optimal structure, represented by point B, has a total mass of 146.9 Kg and an RMS surface distortion of 44.6 microns representing a simultaneous improvement of 52% and 75.4% respectively over the baseline design

2.3 **Modeling and structural dynamics.** IMOS (Integrated Modeling of Optical Systems) program [4] and MSC/NASTRAN software were used to develop finite element models for the CSRL testbed. The models include the primary and the secondary truss structures, the panels, the joints, the fittings, and the actuators. The eigenvalue analysis of the system showed that the lowest natural frequency of the structure is at 10.3 Hz. Figure 3 illustrates the frequency histogram for the first 100 modes of the structure indicating that the dynamic characteristics of the CSRL structure are similar to those of a large flexible structure characterized by low fundamental frequency and high modal density. Figure 4 shows mode shapes representing the first significant primary structure frequencies.

3. CONTROL SYSTEM REQUIREMENTS

The following requirements were developed at CSRL for the control system:

- Line-of-sight (Pointing) accuracy of 2 arc seconds.
- Figure maintenance to within 1 micron (rms distortion) with respect to calibrated surface,
- Control Bandwidth 15-30 Hz.
- Use voice coil actuators to provide an actuator bandwidth of 100-200 Hz.
- Control up to first 20 modes reduce spil lover effect due to neglected modes and dynamics,
- Attenuate vibrations due to gravity, thermal, seismic effects, etc.
- Minimize the "Disturbance" effect of the active control.

4. MATHEMATICAL MODEL

The equation of motion for the system under consideration is given by:

$$M\ddot{q}+Kq=B_1u+B_2f$$

where M is a positive definite symmetric mass matrix, K is a positive semi-definite symmetric stiffness matrix, B_1 is a control influence matrix, B2 is a disturbance influence matrix, and q is the vector of physical coordinates, i.e.. panel displacements. The generalized mass and stiffness matrices are developed via finite element modeling (FEM).

The segmented reflector telescope structure consists of 42 nodes and each node introduces 6 degrees of freedom. Control and performance evaluation of such structure is an exceedingly difficult task. To reduce the number of degrees of freedom and still maintain accuracy, a method called Guyan reduction is used here to reduce the system mass and stiffness matrices by omitting the x and y degrees of freedom of the 42 truss and panel nodes. The reduced mass and stiffness matrices are further used for decentralization and control design.

5. DECENTRALIZATION

Large space structure control design is characterized by high dimensionality and complexity making the controller one of a large order with attendant complexity. Considering that the original state-space model is derived via finite element programming which already uses many assumptions about structural elements, their masses, mass distribution, other properties, and interconnections, etc., the original state space model itself can be assumed to be but an approximation of the real structure. Therefore, instead of corrupting the models further by eliminating more of the available states, the approach of decentralized control appears to be a more realistic way of alleviating the dimensionality problem. The controllers for individual panels are derived using all of the available measurements from the entire structure. Decentralized control is a better approach in terms of designing a realistic controller because it is not based on neglecting available information. Also, in terms of reducing the control spillover effect, decentralized control is better compared to model reduction achieved by drastically reducing the number of states. It is well known that control spillover resulting from neglected modes and dynamics can potentially destabilize a system.

One of the most important aspects of decentralized control is in fault tolerance. In the decentralized approach controllers are designed for individual subsystems while incorporating the interactions among them. In the event of a controller failure, the system can gracefully degrade into one where via proper decentralization, the adjacent panel controllers may compensate for the failed controller. The decentralization could be based on several factors such as time-scale based decentralization, frequency-scale based decentralization or structure-based decentralization. In this paper the natural symmetry of the structure is exploited to decompose the system into smaller subsystems, each of a much lower order. It has merits in reducing the controller size and complexity thereby easing computational burden on the processor. Simpler control algorithms will in turn lead to simpler hardware implementation. The reduced version of the structure is decomposed into six subsystems as follows:

$$x_i = A_{ii}x_i + E_ix$$

where E_i is the subsystem interactions including the effects of disturbances.

The isolated panel components are given by $\dot{x}_i = A_{ii} x_i$, where $A_{ii} = \begin{bmatrix} 0 & I \\ -\omega_i^2 & -2\zeta_i\omega_i \end{bmatrix}$.

6. NEURAL NETWORK BASED CONTROL

Two neural network architectures were developed and tested in simulation for the problem of disturbance rejection of the large segmented space reflector, The development and simulation results of these architectures are presented in the current section. The *first controller* is a novel neural network controller (NNC) whose parameters are adjusted on line [7]. The control algorithm is simple and can be implemented in real time. Unlike other NNCs that are reported in the literature, the proposed neural network controller requires relatively few neurons and its learning algorithm is faster than backpropagation. Stability analysis by a Lyapunov approach is used to determine the convergence properties of the algorithm. The stability is guaranteed with rather mild conditions and certain prior knowledge of the plant to be controlled. The proposed adaptive control consists of a neural network placed in the feedback loop and an adaptation law to adjust the parameters of the net. The *second controller* is a feedforward two-layer neural network. The net work is trained off-line to emulate a dynamic compensator. The training is done by classical back-propagation.

6.1 **Adjustable neural network controller.** The proposed control system is shown in Fig. 5 and is based on part of the doctoral dissertation of J.-A. Luzardo [7] currently in progress. In the following, we describe its main characteristics.

The plant to be controlled is assumed to be *almost strictly positive real (ASPR), i.e.* there exists an output feedback gain matrix K such that the closed loop system $A - BKC$ is *strictly positive real* [6]. It is noted that the value of K *is* not needed for implementation, only its existence is required.

The neural network controller of Fig. 5 is a two-layer network with one hidden layer ($u = N(e)$). The internal network topology is arranged to provide the following outputs (control inputs to the plant):

$$u_i = \sum_{j=1}^{p} \left(\sum_{k=1}^{l} c_{ijk}(t)\sigma(\omega_{ijk} e_j + \theta_{ijk}) \right)$$

where u_i is the *i-th* component of u and e_j is the *j-th* component of $e = y_m - y$, and $\sigma(z) = 1/(1 + e^{-z})$ It is noted that the parameters c_{ijk} are time varying, while the parameters ω_{ijk} and θ_{ijk} are constant. For notational simplicity we denote $\sigma(\omega_{ijk} e_j + \theta_{ijk})$ by σ_{ijk}. Then $u_i = \Sigma_i^T \alpha_i$, where $\Sigma_i = [\sigma_{i11} \cdots \sigma_{i1l} \cdots \sigma_{ipl}]^T$ and $\alpha_i = [c_{i11} \ldots c_{i1l} \cdots c_{ipl}]^T$. Using these definitions, the vector u can be written as

$$u = \begin{bmatrix} \Sigma_1^T & & \\ & \Sigma_2^T & \\ & & \ddots \\ & & & \Sigma_p^T \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix} = \Phi\alpha$$

The adjustable parameters of the neural network controller are adjusted according to the following adaptation law: $\dot{\alpha} = \Gamma_1 \Phi^T e - \Gamma_1 \Gamma_2 \alpha$, where Γ_1 and Γ_2 are two symmetric positive definite matrices chosen according to design criteria. Under these conditions it is proven by Luapunov function arguments that the neural network controller will keep the closed loop system stable and the signals in the closed loop will remain bounded.

6.2 **Simulation results.** The *ASPR* condition in the case of a flexible space structure is guaranteed if the sensors and actuators are collocated and if the measured output is a combination of rate and position measurements [6]. Simulations were performed on the telescope model, under these assumptions, and some of' the simulation results are shown in Fig. 6. The results shown are for a three-panel subsystem, when three sinusoidal disturbances were applied 10 each panel. Only the responses for panel 2 are displayed. The dashed line is the open loop response, whereas the solid line is the response under the neural network controller. It is seen that the neural controller attenuates the response due to disturbance by a factor of 10 or higher. Similar rc.suits were obtained for the remaining panels and for other disturbances with different characteristics.

6.3 **Neural network controller trained by back propagation.** This is a neural controller concept that was investigated as an alternative to the adjustable neural network controller of the previous subsection. As a first step towards the development of this controller,it was decided that the control ler block be trained off-line by standard back propagation, and that its performance be evaluated before proceeding to designing on-line training algorithms. Thus, it was assumed that the system is known, hence there is no need an for on-line neural identifier, and a dynamic compensator was designed to control the known system. The neural controller block was trained off-line to emulate the dynamic compensator, based on input/output data from the compensator. The resulting neural network was placed in the forward loop as a neural implementation of the dynamic compensator, and several simulation studies were performed to evaluate its performance. The simulation tests presented here are for a two-panel subsystem. A sinusoidal disturbance with frequency close to the natural frequency of the simplified model is chosen. Fig. 7 shows the regulation ability of the neural controller, for panel 1. The dashed line represents the positions with the controller off (open loop), whereas the solid line represents the positions with the controller on (closed loop). It is again seen that the controller reduces the disturbance effects by almost a factor of 10.

7. COMPUTER ARCHITECTURE

The drive electronics used in the CSRL testbed is for the purpose of processing the analog output of sensors and interfacing with the segment positioning actuators. The drive electronics is in charge of real-time processing and data acquisition. The computer and graphic setup includes a DSP, a PC and two SUN stations. Figure 8 illustrates the overall computer architecture block diagram. The DSP is the main computational unit and it is responsible for real time control processing, signal generation, and real time directory memory. Access data transfers to a 256 K bytes internal memory block resides on the DSP. The DSP and the SUNS are used to monitor the CSRL experiments via the graphical display of the Kaman sensors reading, the actuator commands, and the mirror segments piston and tilt misalignment. The input/output unit is composed of two 32-channels 16 bit analog to digital and two 18-channel16-bit digital to analog converters.

8. SUMMARY

A lightweight rigid structure has been designed using Pareto optimality technique. The structure exhibits dynamic behavior of a large flexible structure characterized by low fundamental frequency and high modal density. The structure has been decomposed into six subsystems. The control law is designed using the individual subsystems as well as the interconnection properties among them. The control design is performed using neural net works. Performance evaluations are performed utilizing PRO-MATLAB and SIMULINK. The *first neural network controller* adjusts its weights on-line and requires measurements of the positions and velocities, and that the actuators and sensors be collocated. Under these conditions the ASPR condition is satisfied, allowing the NNC to be implemented. The NNC is simple because it adjusts the coefficients of the linear layer (output layer) of the NN. The internal parameters of the sigmoids (hidden layer) are fixed in accordance with the operation range of the system. A special consideration is taken to uniformly distribute the centers of the sigmoids all over the range of operation. The simulation results present satisfactory transient responses and a fast adaptation, although the neural weights were all initialized to zero. As predicted by theory, all the signals remain bounded and all position errors remain small, in the presence of external disturbances. The results from *second neural controller* suggest that a strategy utilizing prior

information about the structure, such as a good model, can be very effective. A dynamic compensator designed from the numerical model is used to train a feedforward neural network. The trained network is shown through simulations to have good disturbance rejection properties and good step following. The weights of the trained network can be used as an initialization point for subsequent on-line weight adaptation algorithms.

9. ACKNOWLEDGMENTS

The research described in this paper was carried out at the Control and Structures Research Laboratory and was supported by the National Aeronautics and Space Administration under grant # NAGW-4103.

10. REFERENCES

[1] Ryaciotaki-Boussalis, Z. Wei, and M. Mirmirani, "Decentralization and PID Controller Design for Large Space Borne Telescopes," IASTED Conference on Modeling and Simulation, Pittsburgh 1995.
[2] Aubrun, Jean-Noel, et al. "Active Control for Segmented Mirror Optical Systems," Research and Development Division, Lockheed Missiles and Space Company, Palo Alto, California.
[3] Hahn, M., Mirmirani, M., Boussalis, H., "Optimal Design of a Truss Structure for a Segmented Reflector," ASME Conference, Boston, 1995.
[4] Stadler, W., Editor, "Multicriteria Optimization in Engineering and in Science," Plenum, NY, 1988.
[5] H. R. Boussalis, C. H.Ih, "Modeling and Stability of Segmented Reflector Telescopes: A Decentralized Approach," Proc 23rd Asilomar Conference, 1989.
[6] Kaufman, H., Bar-Kana, I., and Sobel, K. "Direct Adaptive Control Algorithms." Springer, 1994.
[7] Luzardo, J-A. "Neural Networks for Approximation and Control of Continuous Nonlinear Dynamical Systems.", Ph.D. thesis, Dept. of Mathematics, The Claremont Graduate School, Claremont, CA. 1996,

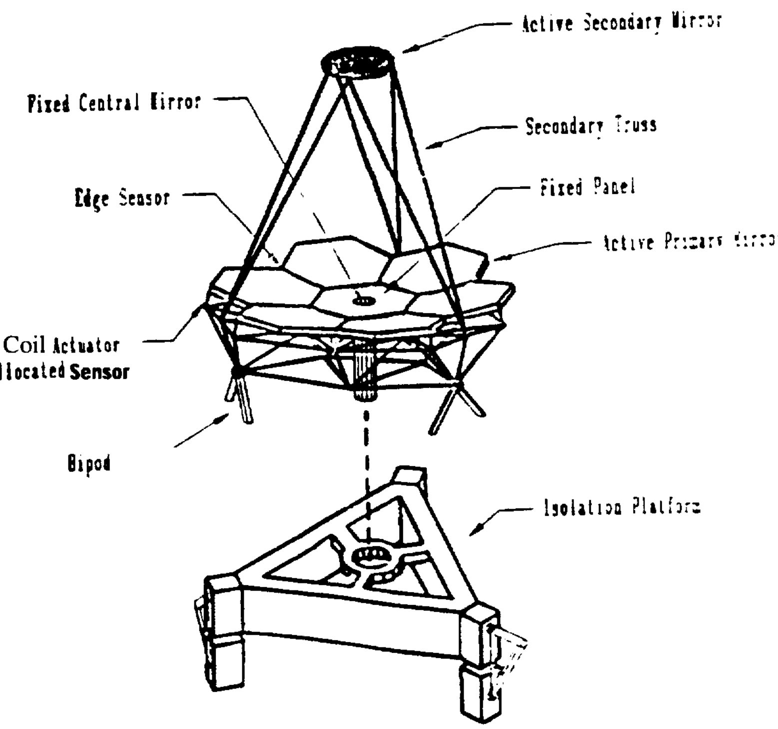


Figure I: The CSRL segmented reflector

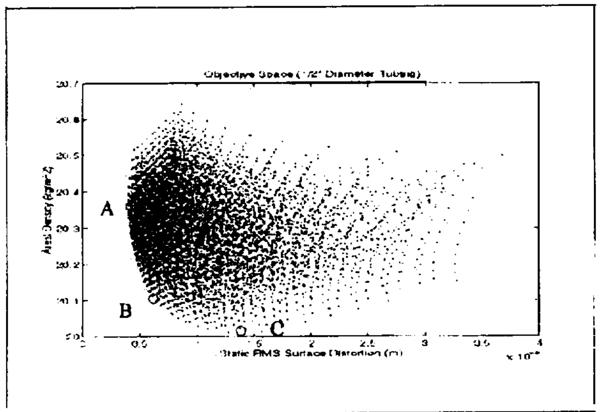


Figure 2: Truss optimization results

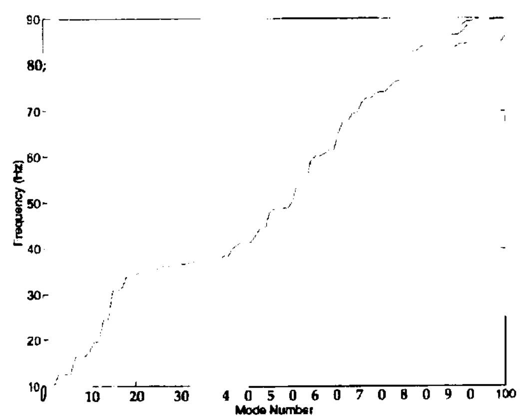


Figure 3: Frequency histogram of the CSRL structure

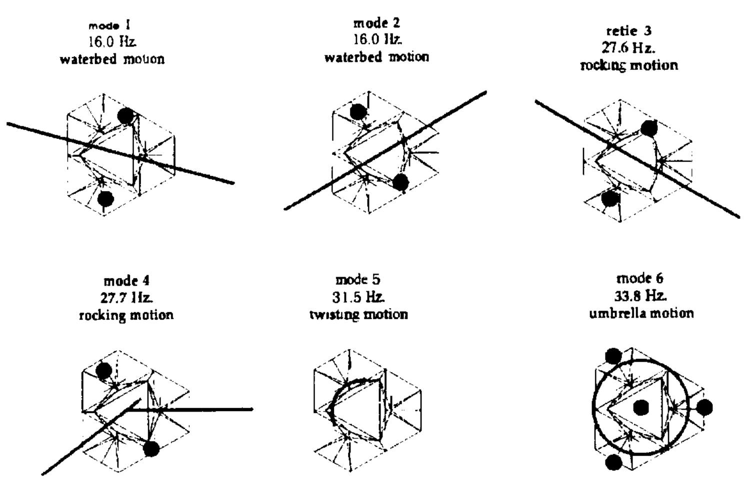


Figure 4: First six mode shapes of the CSRL structure

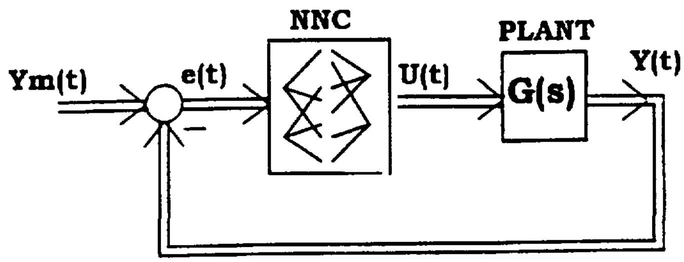
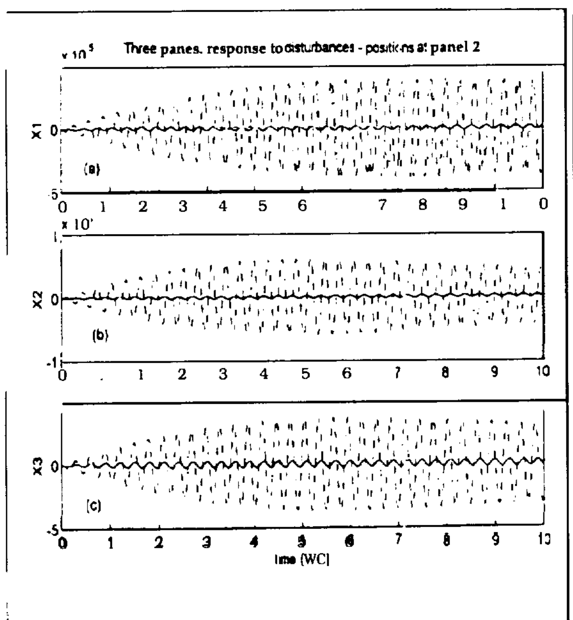


Fig. 5: Block diagram of the neural network controller



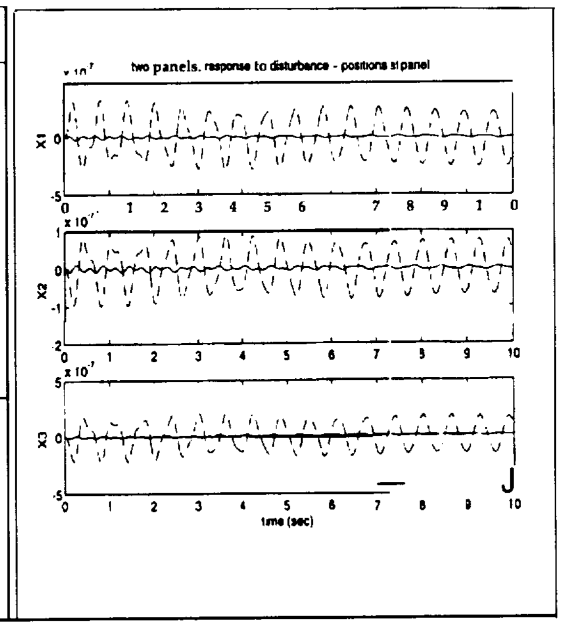


Fig. 6: Adjustable neural network controller. Open loop (dashed line) and closed loop (solid line) responses under sinusoidal disturbances

Fig. 7: Controller trained by back propagation. Open loop (dashed line) and closed loop (solid line) responses under sinusoidal disturbances

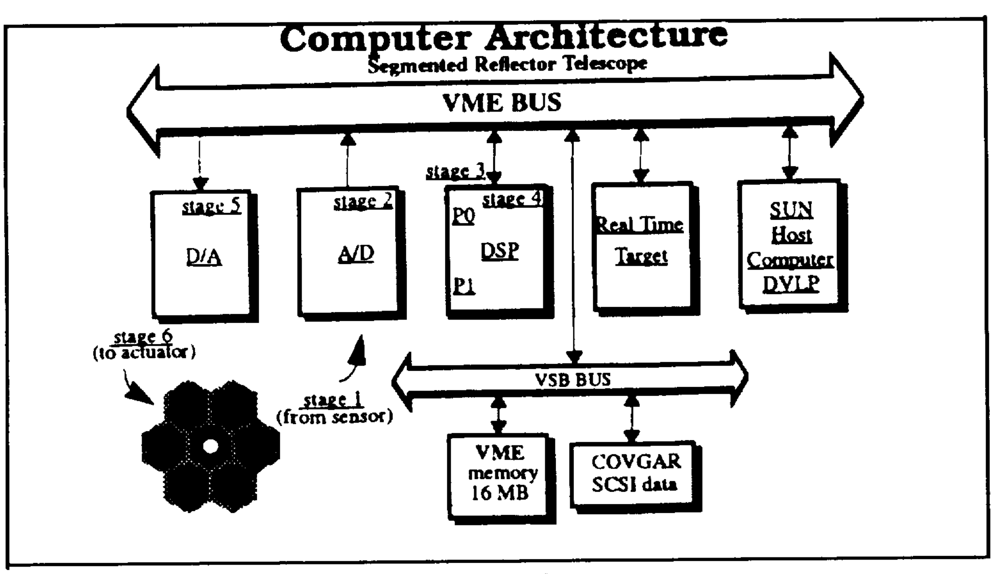


Figure 8

Page intentionally left blank

/ / / / 5 A   5 / / 2 / 6 3

URC97112

A NEW UNIVERSAL ANALOG FUZZIFIER BASED ON OPERATIONAL TRANSCONDUCTANCE AMPLIFIERS

Patricia Saavedra[1] *Jaime Ramírez-Angulo*[2] *and Jorge Zrilic1*[1]

[1]Engineering Department, New Mexico Highlands University, Las Vegas, New Mexico
[2]Klipsch School of Electrical Engineering, New Mexico State University, Las Cruces, NM, 88003

Abstract. A versatile analog **fuzzifier** circuit for membership functions with trapezoidal shape is introduced. The parameters of the trapezoidal function are continuously and independently adjustable. The **fuzzifier** uses operational **transconductance** amplifiers. Experimental results that verify the operation of the **fuzzifier** are presented.

I. INTRODUCTION

Fuzzy logic is an innovative technology to provide engineering systems with human expertise [1]. It is being used in many applications that include industrial automation, process control and data processing. Fuzzification is the basic operation of fuzzy logic. It is used to determine the degree of membership of a system's input and output variables to fuzzy sets. Membership functions are characterized by degree of association curves. One of the most common shapes is the trapezoidal function (Fig. 1). This function includes triangular, S- and Z-shaped functions as special cases. Special purpose fuzzy processor chips have been developed for applications that require high speed. Most fuzzy processors chips are based on digital techniques [2]-[4]. Digital implementations are silicon area intensive, but they have the advantage of easy programmability and easy interfacing to conventional digital systems. Analog approaches are inherently faster and require much smaller silicon area and lower power consumption [5]-[6]. Their main disadvantage is that they are not so easily programmable. In this paper we present a new high-speed current-mode/voltage-mode fuzzifier characterized by a trapezoidal membership function (Fig. 1). The x-position of the comers (Va,Vb, Vc, Vd) of the trapezoidal function as well as its amplitude (Vamp) are independently and continuously adjustable. This feature is specially useful for high speed neuro-fuzzy applications [71.

II. OTA IMPLEMENTATION

The circuit of the proposed analog fuzzifier is shown in Fig. 2b. [t uses four commercial bipolar operational transconductance amplifiers (OTAs) CA3280 which have a Gilbert translinear cell (GTC) between their input terminals for linearization purposes. External resistors R in series with the input terminals are used in conduction with the GTC to to extend the linear input range of the OTA. The output current of a linearized OTA is expressed approximately by $Iout=I_{bias}V_{dif}/(I_{diode}2R)=g_mV_{dif}$ in its linear region. Where V_{dif} is the differential input voltage, I_{diode} and I_{bias} are currents injected to linearizing and bias control terminals of the OTA respectively. g_m is the OTA transconductance gain $g_m=I_{bias}/(I_{diode}2R)$. The OTA control terminals have resistors rd and rb in series. The currentsin these terminals are expressed by: $I_{diode}=(V_{diode}-VSS')/rd$ and $I_{bias}=(V_{bias}-VSS')/rb$ respectively. The voltage VSS'is approximately two diode drops above the negative supply voltage VSS: VSS'=VSS+1.4V. VSS' corresponds approximately to the voltage at the OTA control terminals. I_{bias} determines the output saturation current of the OTA and the input linear range which is defined by $I_{bias}/g_m<V_d<I_{bias}/g_m$.

Silicon junction diodes connected at the output of the active OTAs are used to implement rectifier and catch diodes. Rectification produces unidirectional output currents Io1 and Io2 which are required to generate a trapezoidal membership function with independently adjustable parameters as explained later. Catch diodes are used to prevent large voltage swings at the output of the OTAs which would otherwise lead to poor high frequency performance. Two OTAS (denoted I and II) with equal currents $I_{diode}=I_{bias}$=VDD-VSS' /r are used to set the current I_{diode} of two active OTAs (denoted III and IV) to values IdiodeIII=(Vd-Vc)/2R and IdiodeIV=(Vb-Va)/2R respectively. OTAs III and IV produce output

[1]This work was supported by the NASA/ACE Center for Autonomous Control Engineering and by CIMD/NSF

currents Io2=Iamp(Vin-Vc) / (Vd-Vc) and Io1=Iamp(Vin-Va) / (Vb-Va) with the transconductance characteristics shown in the left and middle traces of Fig. 4a respectively. Equal values are used for the bias currents, $I_{bias}=I_{amp}=V_{amp}-VSS'/r$, of OTAs 111 and IV. This current is used to control the amplitude of the trapezoidal function (I_{amp} is controlled by means of the voltage denoted V_{amp}). An op-amp is used to transform the current Io=Io1-Io2 into the output voltage Vo=IoR. The transfer characteristic of Io has the trapezoidal characteristic of Fig. 1 (or of Fig. 2a).

111. Experimental results

Fig. 3 shows experimental results of the circuit of Fig. 2b. Fig. 3a shows adjustment of the voltages Vb and Vc to transform the trapezoidal characteristic into a triangular one, Fig. 3b shows adjustment of the amplitude with Vamp and Fig. 3c shows adjustment of Voltage V_d to change the slope of the left transition region. Other adjustments are not shown for the sake of space.

IV. Conclusions

A new analog fuzzification scheme was introduced and experimentally verified. It is characterized by a trapezoidal membership function with independently and continuously adjustable parameters. The circuit is very compact and it uses four operational transconductacne amplifiers and an operational amplifer. A BiCMOS integrated version of the same scheme is currently under development and will be reported in a near future. The layout of this circuit is shwon in Fig. 4. This scheme is expected to find application in neuro-fuzzy processors where c adaptation of the membership function para required.

Acknowledgements

The help of Mr. Gerardo Gonzalez for the test of the breadboard prototype of the fuzzifier is acknowledged.

References

[1] RR Yager and L A. Zadeh:" An Introduction to Fuzzy Logic Applications in Intelligent Systems," *Kluwer Academic* pub. 1992

[2] H. Watanabe, W. Dettloff and K. E. Yount, "A VLSI Fuzzy Logic Controller with Reconfigurable, Cascadable Architecture: *IEEE JSSC*, pp. 376-382, v. 25, No. 2, 1990.

[3] WARP, Fuzzy Logic Processor, SGS-Thompson

[4] NLX220, Fuzzy Logic Controller, American Neuralogix Inc.

[5] T. Miki, H. Matsumoto, K. Otho and T. Yamakawa, "Silicon Implementation for a Novel High-Speed Fuzzy Inference Engine Mega Flips Analog Fuzzy Processor," *Journal of Intelligent and Fuzzy Systems*, Vol. 1, No. 1, pp. 27-42,1993

[6] "A BiCMOS Universal Membership Function Circuit," J. Ramirez-Angulo, 1995 *IEEE Int. Symp. on Circts. and Syst.*, Seattle, WA, April 30, May 3,1994, pp.275-278

[7] B. Kosko, "Neural Networks and Fuzzy Systems, Prentice Hall, 1992

[8] "Analogue IC design, The current mode approach,", Chapter 6, by B. Gilbert, edited by C. Tomazou, F.J. Lidey and D.G. Haigh, Peter Peregrinus, LTD, 1990, London.

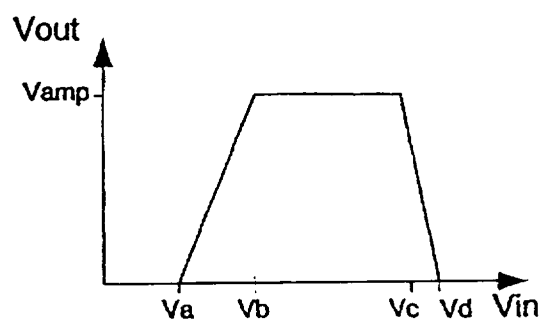


Fig. 1 Trapezoidal membership function

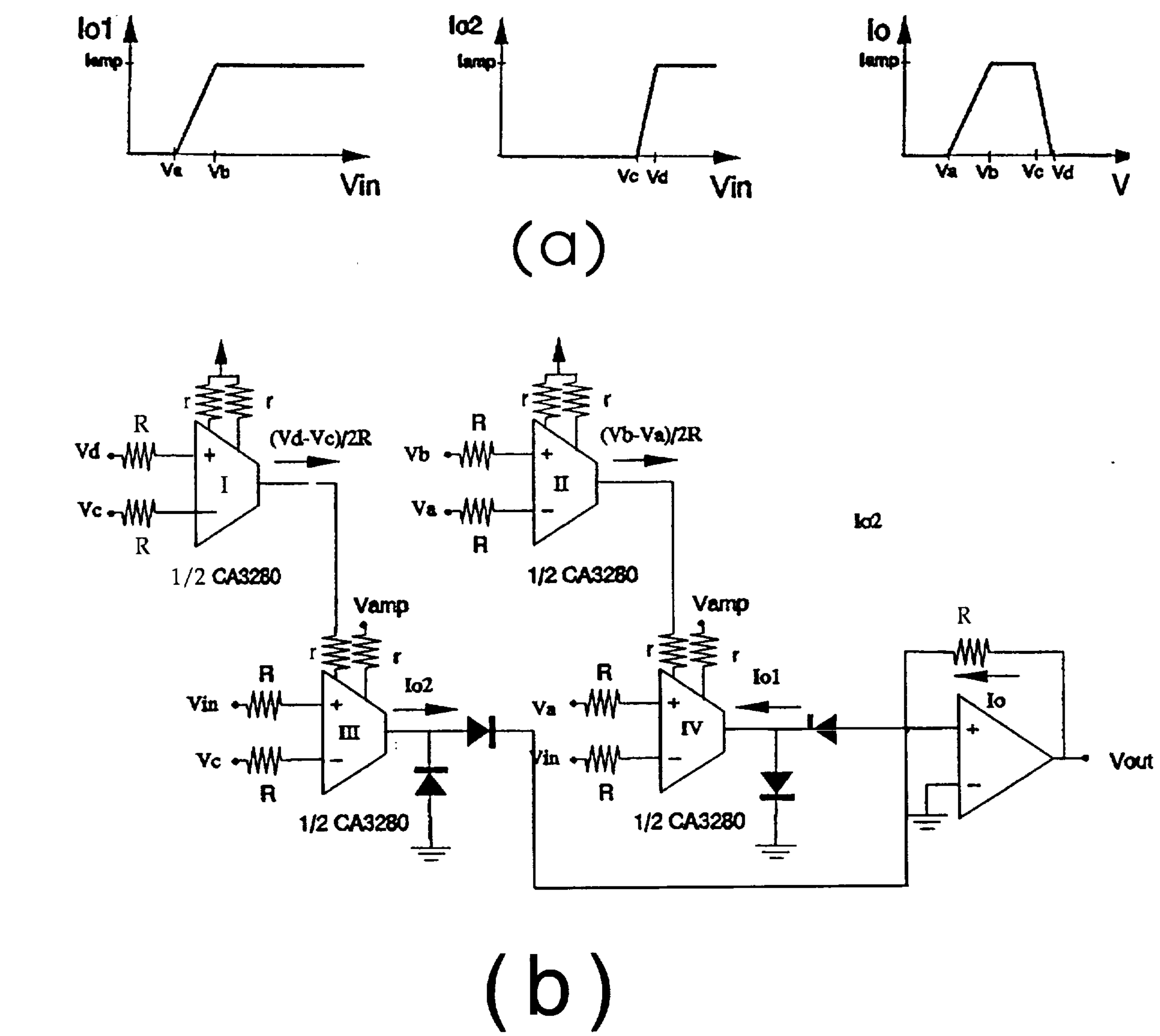
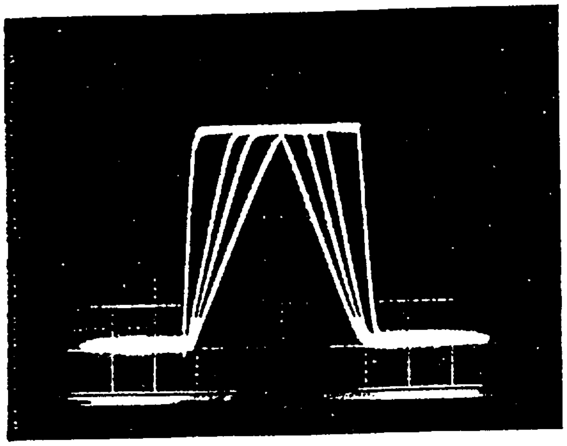
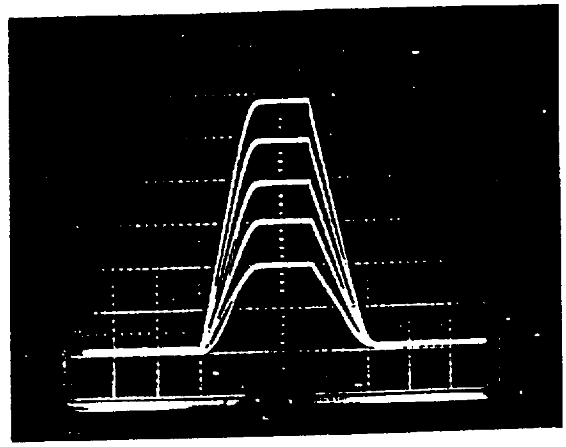


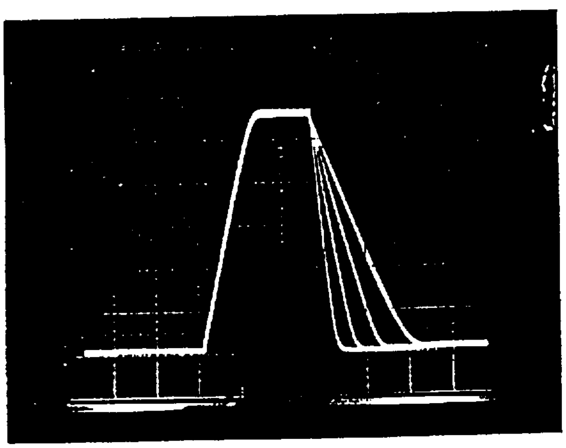
Fig. 2 (a) Generation of trapezoid characteritic (b) Implementation of fuzzifier using commercial OTAs





(b)

(a)



(c)

Fig. 3 Experimental results (Vertical and Horizontal axis 0.5V/div): (a) Vb and Vc adjustment, (b) amplitude adjustment, (c) Vd adjustment

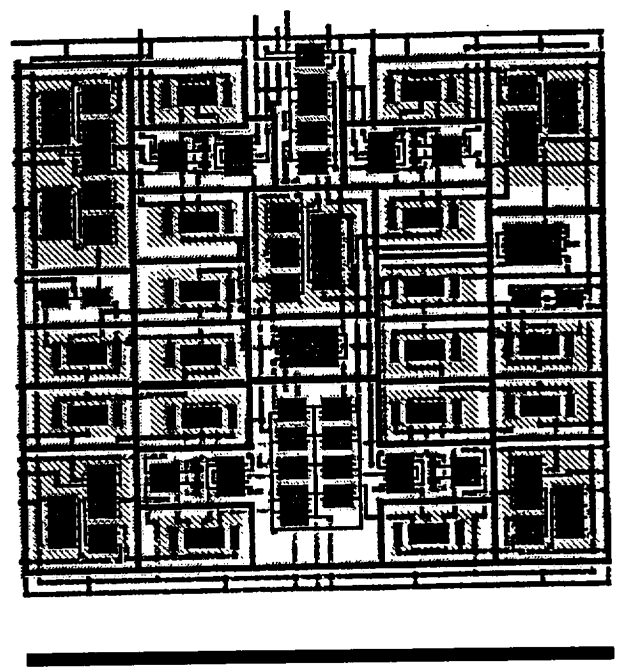


Fig. 4 Layout BiCMOS integrated version of analog fuzzifier currently under fabrication. Reference line on bottom measures 240μm.

ST13/51

URC97113

CODE - An Environment for Multi-Developer Software Development

Raikanta Sahu, Malcolm J. Panthaki, and Dr. Walter H. Gerstle
Department of Civil Engineering
University of New Mexico
Albuquerque, NM 87131

Abstract

This paper describes CODE (CoMe T Development Environment), an environment for a *multi-developer* software *development effort taking place in the Civil Engineering Department at the University of New Mexico. CODE is a collection of guidelines,* standards, and tools *to facilitate* development *of* a large software *system* called *CoMe T (Computational Mechanics Toolkit). A* large software system mandates *strict discipline in the* various *stages of software development. The* software *becomes especially susceptible to* corruption and *misinterpretation* in a multi-developer *environment. Identifying* and maintaining *consistency in* every *aspect of the software development process goes* a *long* way *towards minimizing the above stated problems. We describe the* guidelines we *have adopted to* achieve consistency. *Another important aspect of* software *development is maintaining* versions *of files* and *modules. However, the capability to* merely *maintain* versions *of files and modules proved to be less than adequate for our purposes. We have developed* a *set of requirements for CODE which we think are appropriate to multi-developer* software *development* environments involving large projects. *Dictated by our requirements we have developed tools that extend the capabilities of a public domain* version control *software* package called *CVS (Concurrent* Versions *System). Tools in CODE take advantage of our coding* standards *to automatically* generate *documentation. The tools of CODE have been* designed *to be portable; they depend on CVS, POSIX, and shell scripts.* Currently *they* run *in* a *UNIX* environment and use *the World Wide Web for* making *CoMe T's* documentation *automatically available to the developers. CODE can be used, with* minor *modifications, for any large,* multi-developer software *development effort.*

1. Introduction

Figure 1 depicts the nature of the software development process (SDP) [Schach]. The main goal of a software development environment (SD E) is to provide the tools and infrastructure to ease every step of the SDP.

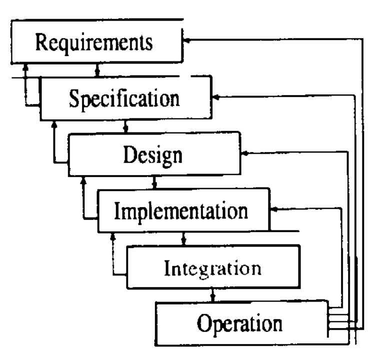


Figure 1. The software development process

The nature of the tools and infrastructure of an SDE must be such that they facilitate:

- Creation and publication of documentation
- Management of changes to documentation and source code

- Building the software from source code
- Unit and integration testing

In most cases, software houses have developed their own methods and tools for documenting. Tim-e are numerous configuration management software packages to help with the management of changes to documentation and source code [Eaton], To ease the task of building the software there is UNIX make, clones of UNIX make, and similar products, There are also integrated development environments that let editing, compiling, and linking to take place in one environment. Most people have their own set of problems and tools to test their software; this is largely undocumented.

A large software system mandates strict discipline in the various stages of software development. The software becomes especially susceptible to corruption and misinterpretation in a multi-developer environment. Identifying and maintaining consistency in every aspect of the software development process goes a long way towards minimizing the above stated problems. An excerpt from a recent publication [Maring] speaks volumes about consistency:

> When application developers use multiple techniques or facilities to design anti build solutions, an element of individual judgement invariably emerges.
>
> ...
>
> Thus, idiosyncrasies are introduced into the design that detract from reusability, make debugging and enhancement difficult, and complicate maintenance.

A large software system called CoMeT (Computational Mechanics Toolkit) is being developed in the Civil Engineering Department at the University of New Mexico. CoMeT is designed using an object-oriented approach and implemented using the C++ and Scheme languages. This is a multi-developer effort involving graduate students from Civil Engineering and Computer Science. We needed a set of guidelines anti standards to maintain consistency in our **work.** We also needed a set of tools and infrastructure to help us with the SDP. The result of our quest for an SDE lead us to what we call CODE (CoMeT Development Environment). CODE is the set of tools and standards, some new and some old, that constitute the SDE for CoMeT.

2. Elements of CODE

2.1. Standards for documentation

The main elements of documentation are text and figures. With the availability of browsers for documents conforming to the HTML (Hyper Text Markup Language) standard, it was an easy decision for us to adopt it for writing the text oft he requirements, specification, and design documents. The popular HTML browsers, like netscape and mosaic, support display of images in the GIF format. That was reason enough for us to maintain our figures in the GIF format. Files that define the interface to the software are maintained in plain text format. We decided to make our documents available to the developers using the World Wide Web.

2.2. Standards for coding

Writing software that is easily read, debugged, and maintained requires a different outlook than writing programs for homework. Past experiences of developers (in both C and C++) have made it possible to identify code writing practices that are more robust than others [Lindal, Henricson].

Writing robust code requires guidelines for physical layout of code and following some provisions of the language while avoiding others. [n addition, writing extensive comments within code makes for readable and understandable code.

2.2.1. Guidelines for physical layout

The important elements of physical layout are:

- File header - In the file header, we give general information about the file such as the name of the file, its purpose, its creator, its creation date, and its current owner. There is also a slot to provide any special notes that may be of interest to users and developers.

- Function header - A function header appears before the implementation of every function. This header tells the purpose of the function, its creator, and its date of creation. There is also a slot to provide any special notes that may be of interest to developers.
- Block construction - Most programming languages provide features to pm-form tasks repeatedly in a loop. Indenting the code inside the loop makes it more readable. However, if the amount of indentation is not constant, the code becomes difficult to read. We decided to be disciplined about indenting and have been indenting code in a block by exactly 3 spaces everywhere.
- Layout of expressions - We decided to provide spaces around the +, −, and = operators and to specifically not provide spaces around the * and / operators. This has made the expressions in our code very readable.

2.2.2. Guidelines for robust code

We have identified a set of about 20 guidelines for writing robust code. Some of the important ones are:

- Prevent multiple inclusion of files
- Make identity of globally accessible objects distinct
- Never specify member data in the `public` or protected section of a class
- Declare the member functions not modifying the state of an object to be `const`
- Provide copy constructors and define assignment functions for classes that :~llocatf?/(deallocate memory
- Make the destructor virtual for all base classes that have a virtual function
- Do not return non-const reference or pointer to member data through public member functions of a class
- Never return a reference or pointer to a local variable in a public function
- Do not use explicit casts

2.2.3. Guidelines for commenting code

We have identified consistent ways of commenting almost ever-y aspect of writing code. The most important comments deal with the interface of global functions and class member functions. An instance of commenting a member function is shown below.

```
void get_parents(EntityType type, DLCDDDAGNodeList& resultNodeSet);
//R void
//I type
//I- The type of the parents to be returned.
//O resultNodeSet
//O- A reference to a set into which the results of the query will be
//O- deposited.
//- This function returns the parents with the given type for the current
//- node.
```

We have written scripts that take advantage of our commenting guidelines to automatically generate documentation for interactive Scheme commands and make them available for browsing through the World Wide Web.

2.2.4. Naming conventions

We have identified consistent ways of naming classes, files, variables, and other objects. For example, the names of classes indicate where they fit in the architecture, the names of all the pointer variables end with `Ptr`, the names of all list variables end with `list`, etc. The naming (conventions clearly indicate the intent of the object. When using the naming conventions becomes a habit, writing code also becomes easier and faster.

2.3. Build

Building software requires keeping track of dependencies. The UNIX make utility provides the basic functionalities to perform simple builds. We decided to use the freely available GNU make since it provides some addition al helpfu l features.

One of the major deficiencies of the make utilities is their lack of support for automatical - generating dependency databases whenever a target is updated. (,ener-sting dependency databases is specially important. when using the compiler or the linker. In languages that allow the inclusion of one file in another, one can make a distinction between the physical and the logical contents of an including file. For example, if file A. C includes file B. h, any changes to file B. h changes the logical contents of file A. C, even though physical contents of the latter may not have changed. Now, let's assume that A. o depends on A. C. When B. h is changed, A. o should be considered out of date. This becomes more important in a multi-developer environment where mutually dependent files need to be modified by developers without worrying whether their mollifications will cause any major problems to others.

We created tools to maintain dependency databases that keep track of the dependencies of both global and local files. The make program uses these databases to bring all target files up to date whenever any included files are modified.

2.4. Configuration Management

Configuration management is the most important part of an SDE. Our key goal was to ensure consistency in the repository. At any time, the latest versions of the code in the repository ought to be compilable and linkable. We started off using CVS (Concurrent Versions System), a package available in the public domain [CVS]. We quickly realized that CVS was designed keeping in mind a development philosophy different from ours. In order to make CVS work for us, with our development philosophy, we had to write tools that do substantially more work before using CVS to do the raw version management task.

CVS maintains a repository of all the files on which it keeps versioning information. Users may check out files they want to modify. After they are satisfied with the modifications, they may wish to check in the modified files. This basic concept is extended to directories and to modules that can consist of multiple directories. When developers want to compile and link in order to test their modifications, they must checkout all the files into their work area. In a multi-developer environment, changes made by one developer must be incorporated by every other developer **in his/her** work area. This will require recompiling file they never modified. In a large project, like CoMeT, the additional time taken for recompiling is expensive and bothersome.

We decided that a user should not keep files s/he is not modifying in her/his work area. The distribution of various files during development is shown in Figure 2. When a developer checks in a file, we used CVS' trigger mechanisms to make a cleartext copy of the file in an appropriate place and remove the file from the developer's work area.

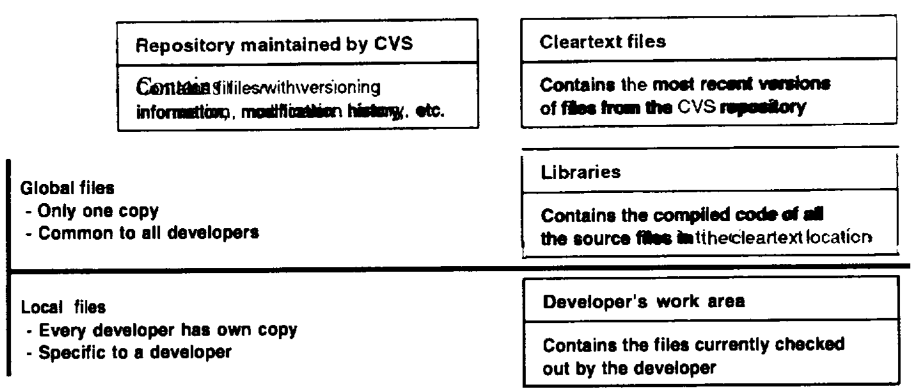


Figure 2. Distribution of files during development

When an implementation file is checked in, we made sure to compile it and put the object code into a library. That way, developers could link their own small set of object code files with the global libraries to

make an executable. This method worked quite well for some time until we ran into global libraries that did not represent the interface properly. Figure 3 shows how dependency among files creates the potential for corruption of global libraries.

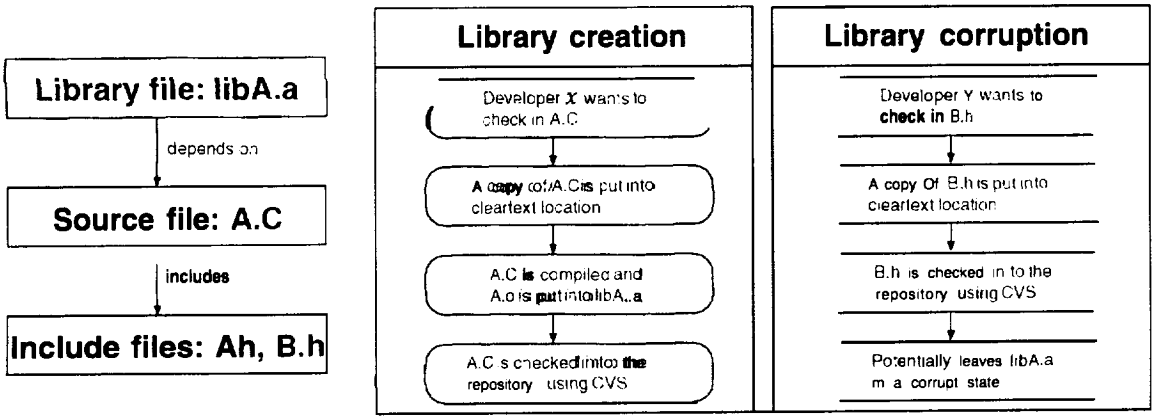


Figure 3. Inadvertent corruption of libraries

To maintain a set of libraries that are always up to date, a modified file must not be put into the repository if it can introduce inconsistencies among interface, implementation, and usage. We *came* up with the algorithm loosely described in Figure 4 to meet our requirements.

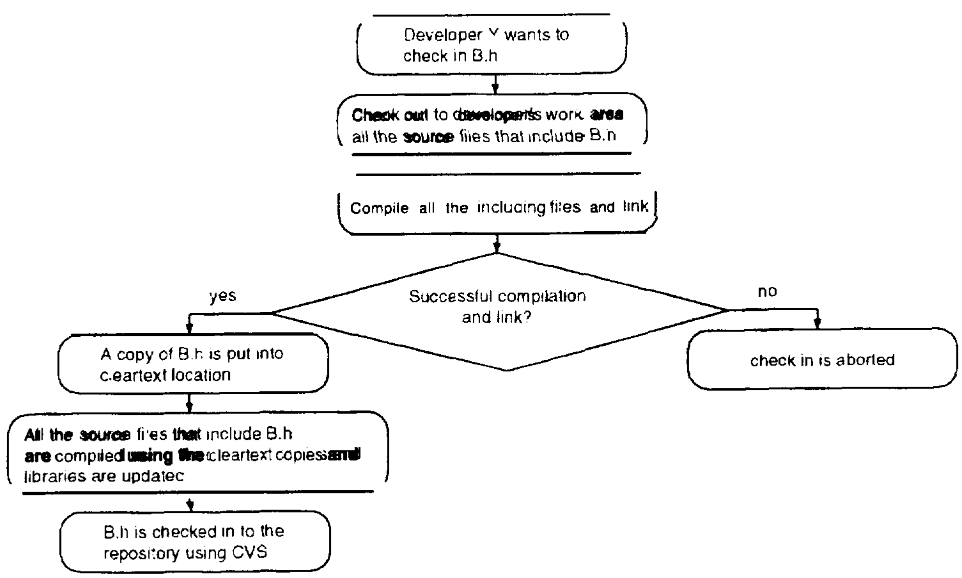


Figure 4. Algorithm to maintain up to date libraries

One of the important tools of CoDE takes care of keeping the global libraries in a clean state by following the algorithm shown in Figure 4. It uses POSIX [Lewine] for most of the work before using CVS for the raw check in. The libraries always represent the state of the most recent versions of the files in the repository. As a result, the repository is always in a clean state. At any time, a developer may check out the entire repository with the assurance that all the source files are Guaranteed to compile successfully and the resulting object files are guaranteed to link successfully. Maintaining a clean repository and global libraries required some extra book keeping. We needed to be able to generate the reverse of the dependency database, i.e., given an interface file, we needed to be able to determine the set of files that use it. We needed to keep track of the location of all files with respect to the root of the repository. We needed to maintain global and

developer specific dependency databases. The dependency database of the developer is always given higher priority than the global database.

After a file is checked in to the repository, the tools of CoDE remove all traces of the file from the user's work area. They also send out email messages to the developers whenever a file is checked in. The changes/comments logged by the developer are appended to the email messages.

2.5. Testing and debugging

We have identified a consistent method for performing unit. tests, i.e. testing of the implementation of single classes or classes closely related to one another. We have also come up with a mechanism to let, developers debug portions of the code they are responsible for. By using different flags, a developer can turn on/off messaging switches at run time to aid in debugging,

3. Concluding Remarks

We described the elements of a multi-developer software development environment called CoDE. CoDE is an aggregate of tools and standards that are currently being used for development of a multi-developer software system called CoMeT. The most important, features of CoDE are:

- It always maintains the most Up to date documents on the World Wide Web.
- It always maintains the repository in such a state that all the source files can be compiled and linked successfully.
- It maintains a set of libraries that are always consistent with the latest versions of' files in the repository.
- It minimizes compiling by individual developers by removing files from their work area when they are checked in to the repository.

The tools of CoDE use CVS, POSIX, and shell scripts. They can be ported to any computer environment that supports the above. Currently they run in a UNIX environment.

The requirements of CODE are appropriate to any large, multi-developer software development effort. CODE can be used, with minor modifications, for any such project.

4. Acknowledgement

We are grateful to the Albuquerque Resource Center of the High Performance Computing Education and Research Center at the University of New Mexico for the financial support extended to Raikanta Sahu and making available their computing resources during this work. We also thank the developers of CoMeT for providing input during the development of CODE.

References

[CVS] URL: ftp: //prep .ai.mit.edu/pub /gnu/cvs-1.9.tar.gz

[Eaton] Eaton, D.: *Configuration Management Tools Summary*, URL: http://www.iac.honeywell.com/Pub/Tech/CM/CMTools.html

[Henricson] Henricson, M. and Nyquist, E.: *Programming in C++, Rules and Recommendations*, URL: http://www.arc.unm.edu/~rsahu/ellemtel.ps

[Lindal] Lindal, J.: *Software* Construction URL: http: //www.arc.unm.edu/~rsahu/software.ps

[Lewine] Lewine, D.: *POSIX* Programmer's Guide, O'Reilly & Associates, Inc., 1994.

[Maring] Maring, B.: *Object- Oriented Development of Large Applications*, IEEE Software, pp 33-40, May 1996.

[Schach] Schach, S. R.: *Practical Software Engineering*. Aksen Associates Incorporate Publishers, Homewood, IL, Boston, MA, 1992.

/// / ' 3/4 /3

URC97114

Land Management in the Tropics and its Effects on the Global Environment: the NASA Institutional Research Award

Schaefer', D.A., T.M. Aide[2], J.D. Chinea[3], N. Fetcher[2], M. Keller[4], S. Molina[5], J.A. Molinelli[6], J.R. Thomlinson[1], G.A. Toranzos[2], R.B. Waide[1], B.R. Weiner', J.K. Zimmerman]'and X. Zou[1].

'Institute for Tropical Ecosytem Studies, University of Puerto Rico, Rio Piedras, PR 00936. 'Department of Biology, University of Puerto Rico, Rio Piedras, PR 00931. [3]Department of Forestry, University of Illinois, Urbana, IL 61801. [4]International Institute of Tropical Forestry, Rio Piedras, PR 00928. 'Department of Biology, Catholic University of Puerto Rico Ponce, PR 00732. [6]Environmental Studies Program, University of Puerto Rico, Rio Piedras, PR 00936. 'Department of Chemistry, University of Puerto Rico, Rio Piedras, PR 00931.

Introduction

Dramatic-land use changes currently underway in the tropics are having important biological and environmental effects, not all of which are well understood. Many of these effects can be grouped as follows:

How do land use changes influence water supplies and flood severity? Replacing forests with cropland or grassland increases runoff, while afforestation has the opposite effect (Bosch and Hewlett 1982, Bruijnzeel 1990).

Removal of forest cover increases soil erosion (Doughs 1967, Sanchez 1976, Lal 1981, Baharuddin 1988), an effect largely attributable to mechanical disturbance of the soil surface during tree removal, and to reductions of soil surface cover by understory vegetation and plant litter (Bruijnzeel 1990). The resulting sediment and dissolved chemical exports can reduce water reservoir storage capacity (Wooldridge 1986), reduce soil productivity (Bruijnzeel and Wiersum 1985), alter the chemistry of streams (Uhl and Jordan 1984), and reduce algal and coral reef productivity below stream out lets (Peterson and Festa 1984, Cloern 1987, Miller and Cruise 1994).

Soil physical, chemical, and microbial properties can be profoundly altered by reduction of forest cover, and in most cases these changes lead to reduced soil fertility (Brinkmann and Nascimento 1973, Dias and Nortcliff 1985, Gillman et al. 1985, Alegre et al. 1986, Lugo and Sanchez 1986, Buschbacher et al. 1988, Rodriguez et al. 1993).

Plant and animal community composition, biodiversity, and successional processes are all altered when lands are cleared, and fill community recovery with reforestation may depend on several spatial processes (Opler et al. 1977, Uhl et al. 1988, 1990).

Removal of forest cover alters water and energy balances, with potential effects on regional climates (Salati et al. 1979, Shukla et al. 1990, Henderson-Sellers et al. 1993). Increased releases of CO, and N_2O from soils following forest removal may alter the thermal balance of the troposphere, with implications for global warming (Duxbury et al. 1993, Keller et al. 1993, Keller et al. 1996 a, b).

In each of these areas, substantial research has been completed or is currently being performed. However, our understanding could be substantially improved by making repeated measurements over long periods of time in tropical areas undergoing land-use changes. Since land-use changes are happening rapidly and are already so extensive, we may not have the luxury of waiting for results from such long-term studies. Comparing measurements made at similar sites that have known climates, soils, and land-use histories is an alternative that can more rapidly provide information relevant to land-management decisions and their implications for global processes..

Another important factor is the impending launch of NASA's Mission to Planet Earth satellites over the next few years. As land-use changes in the tropics are remotely sensed with this new technology, certain physical, chemical, and biological consequences can be anticipated, and quantified as they occur.

The NASA Institutional Research Award (NASAIRA) is a five-year project designed to utilize the unique history of land-use changes in Puerto Rico and the spatial data sets already available (both outlined below), as well as current and subsequent remotely sensed data, to address biological and environmental effects of land-use changes in the tropics.

Over the last 60 years much of the land area in Puerto Rico has been abandoned from agriculture, and secondary forests have been allowed to regenerate naturally (Figure 1). Aerial photography of the entire island has been obtained repeatedly during that period. Detailed land-use classifications are being prepared through time by interpretation of those aerial photographs, so that sites can be identified where particular agricultural practices ended at particular times.

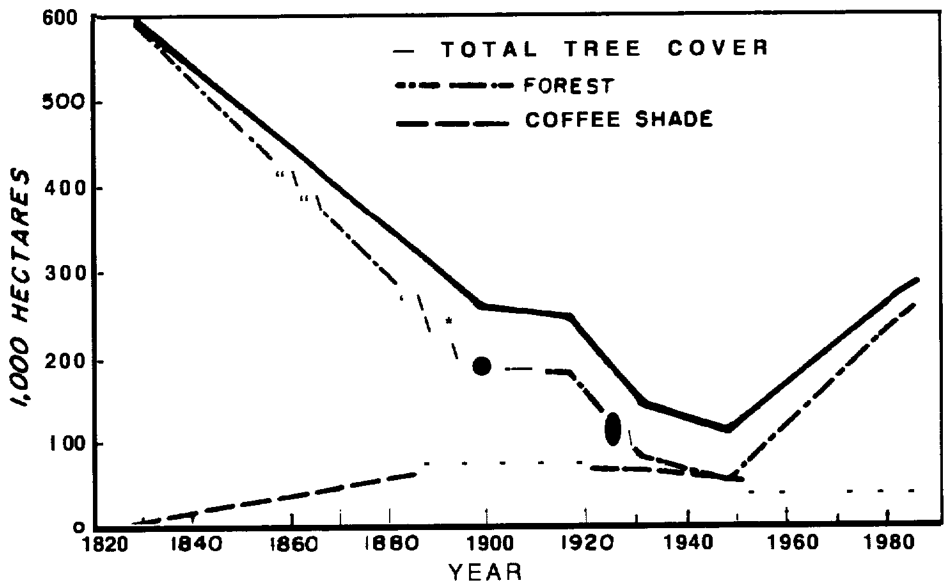


Figure 1. Trends of forested area in Puerto Rico, 1830-1985. The minimum value (in the 1940's) represents approximately 7% of the island area. After Birdsey and Weaver (1987),

Several additional spatially explicit data sets for the island are also available, or are now being prepared (Table 1.). With this information, sites with similar soil and environmental conditions where particular agricultural practices ended at different times are being identified. These groups of sites with similar starting conditions, but with different land uses or times since agricultural abandonment, represent chronosequences of forest recovery. Within these chronosequences, changes in vegetation communities, plant physiology, soil physical, chemical, and biological characteristics, and biogeochemical fluxes are being compared through time.

Table 1. Spatially-explicit data currently available or being developed for Puerto Rico within the NASAIRA project.

AOCI, AVHRR, MSS, SPOT, TIMS, TM, and TMS/CAMS imagery; various dates since 1983.
Digital line graphs of elevation, hydrography, transportation, and political boundaries.
Surficial geology.
Soil associations, series, and related physical and chemical information.
Spatial models of precipitation and temperature.
Daily flow records and periodic water quality sampling for streams.
Holdridge life zones.
Island-wide land-cover types in 1977.
Land use for specific sites in 1936, 1962, 1971, 1977, 1983, 1988, and 1993.

Daily streamflow monitoring by the U.S. Geological Survey (USGS) began in 1943. The number of monitored streams increased to 5 in the mid- 1950's, 25 in the mid- 1960's, 45 in the mid- 1970's, 50 in the mid-1980's, to 70 at present (Figure 2.). Currently monitored streams drain approximately 75% of the island area. In addition, water quality parameters (sediment transport, chemistry, and microbiology) are monitored periodically by the USGS on streams that now number approximately 80. Together, these databases represent the greatest intensity of hydrological and hydrochemical monitoring for any area in the tropics. These data enable us to develop and test several hypotheses concerning how land-use changes influence watershed functions.

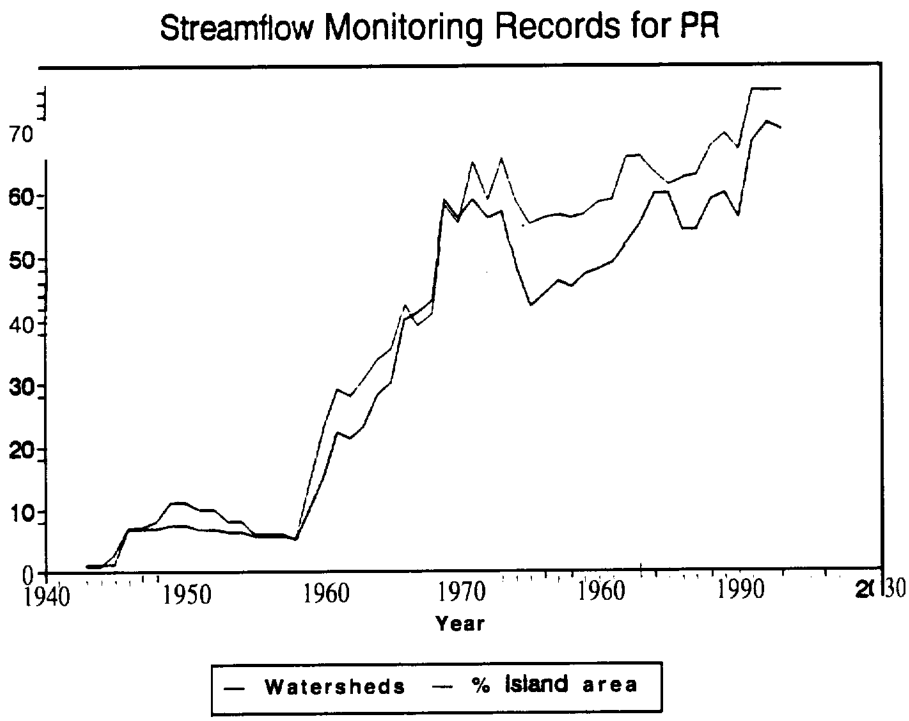


Figure 2. The number of USGS stations recording daily streamflows and the percentage of the island of Puerto Rico covered by those watersheds, 1940-1994.

Factors Monitored in the Chronosequences

Plot studies are used to examine forest recovery pathways and rates under different kind-use histories. They also examine barriers to secondary forest succession, such as plant community types, distance to seed sources, biological vectors, and environmental factors. Physiological characteristics of recovering forests are being explored, including photosynthesis, water relations, canopy spectral reflectance properties, and foliar chemistry. In the same plots we examine soil physical, chemical, and biological characteristics, nitrogen transformations in soils, and soil microbial and termite emissions of greenhouse gases.

Watershed studies are utilized to integrate biological functions at larger scales. Stream channel erosion and sediment movements are examined. Water, chemical, and sediment fluxes and microbial quality of streams are monitored, based on the substantial USGS data base and additional sampling. Watershed hydrology is based on spatial models of rainfall inputs and evapotranspiration in addition to the streamflow data. Finally, atmospheric studies involve laser spectroscopic examination of chemical kinetics of the troposphere.

NASAIRA Products

In the NASAIRA we are developing spatially explicit representations of the land-use history and several other environmental factors in Puerto Rico. We are also developing several predictive models in certain land-use chronosequences, and testing those models in an independent set of sites. We will also test the applicability of the project results in the Dominican Republic, which is now extensively deforested, much as Puerto Rico was 50 to 60 years ago.

Those models relate to forest succession following different agriculture types, forest carbon and water balance, relationships between forest characteristics and remotely sensed data, relationships between land-use changes and watershed hydrology, chemical and sediment fluxes, soil characteristics, and greenhouse gas fluxes. We are also developing an improved understanding of chemical reactions in the troposphere.

Some preliminary results of this project have already been published or are now in press, including studies of forest successional processes (Aide et al. 1995, 1996, Thomlinson et al. 1996), soil fauna (Gonzales et al. 1996, Zou and Gonzales 1996), soil bacteria (Rodriguez et al. 1993), and soil greenhouse gas emissions (Keller et al. 1996a, b).

Conclusions

The NASAIRA project in Puerto Rico is designed to take full advantage of the unique land-use history of this island, as well as several spatially explicit databases. To be sure, the full range of implications of land-management changes in the tropics will not be addressed by the NASAIRA project in Puerto Rico. However, this observational and modeling framework offers an important opportunity for combining unusually complete records of historical land uses and environmental factors with ground-based observations and remotely sensed data to predict environmental changes associated with land-use changes in the tropics.

As data from NASA's Mission to Planet Earth becomes available, our NASAIRA products will increase in value. Certainly, deforestation and other land-use changes will continue in the tropics. We intend that the results from this project will be used to reduce negative environmental impacts of these changes, and to aid in making land-management decisions in these regions.

Literature cited

Aide, T. M., J.K. Zimmerman, L. Herrera, M. Rosario, and M. Serrano. 1995. Forest recovery in abandoned tropical pastures in Puerto Rico. Forest Ecology and Management 77: 77-86..

Aide, T.M., J.K. Zimmerman, M. Rosario, and H. Marcano. 1996. Forest recovery in abandoned cattle pastures along an elevational gradient in northeastern Puerto Rico. Biotropica (in press).

Alegre, J.C., D.K. Cassel, and D.E. Bandy. 1986. Effects of land clearing and subsequent management on soil physical properties. Soil Science Society of America Journal 50:1379-1384.

Baharuddin, K. 1988. Effect of logging on sediment yield in a hill dipterocarp forest in Peninsular Malaysia. Journal of Tropical Forest Science 1:56-66.

Birdsey, R.A. and P.L. Weaver. 1987. Forest area trends in Puerto Rico. U.S. Forest Service Research Note SO-33 1. Southern Forest Experiment Station, New Orleans, LA.

Bosch. J.M. and J.D. Hewlett. 1982. A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration. Journal of Hydrology 55:3-23.

Brinkmann, W.L.F. and J.C. de Nascimento. 1973. The effect of slash and bum agriculture on plant nutrients in the Tertiary region of central Amazonia. Turrialba 23:284-290.

Bruijnzeel, L.A. 1990. Hydrology of Moist Tropical Forests and Effects of Conversion: A State of Knowledge Review. UNESCO International Hydrology Program, Paris.

Bruijnzeel, L.A. and K.F. Wiersum. 1985. A nutrient balance sheet for *Agathis dammara* Warb. plantation forest under various management conditions in central Java, Indonesia. Forest Ecology and Management 10:195-208.

Buschbacher, R. J., C. Uhl, and E.A.S. Serrao. 1988. Abandoned pastures in eastern Amazonia [1. Nutrient stocks in the soils and vegetation. Journal of Ecology 76:682-699.

Cloern, J.E. 1987. Turbidity as a control on phytoplankton biomass and productivity in estuaries. Continental Shelf Research 7:1367-1381.

Dias, A. C.C.P. and S. Nortcliff. 1985. Effects of two land clearing methods on the physical properties of an Oxisol in the Brazilian Amazon. Tropical Agriculture (Trinidad) 62:202-212.

Douglas, I. 1967. Natural and man-made erosion in the humid tropics of Australia, Malaysia, and Singapore. International Association of Hydrological Sciences Publication 75:17-30.

Duxbury, J.M., L.A. Harper, and A.R. Mosier. 1993. Contributions of agroecosystems to global change. In Agricultural Ecosystems Effects on Trace Gases and Global Climate Change, L.A. Harper, A.R. Mosier, J.M. Duxbury, and D.E. Rolson (eds.), Agronomic Society of America Special Publication No. 55, Madison, WI.

Gillman, G.P., D.F. Sinclair, R. Knowlton, and M.G. Keys. 1985. The effect on some soil chemistry properties of the selective logging of a north Queensland rainforest. Forest Ecology and Management 12: 195-214.

Gonzalez, G., X. Zou, and S. Borges. 1996. Earthworm abundance and species richness in rehabilitated agricultural fields: Effects of tree plantations. Pedobiologia (in press).

Henderson-Sellers, A., R.E. Dickinson, T.B. Turbridge, P.J. Kennedy, K. McGuffie, and A.J. Pitman. 1993. Tropical deforestation: modeling local- to regional-scale climate change, Journal of Geophysical Research 98: 7289-7315.

Keller, M., E. Veldkamp, A.M. Weitz, and W.A. Reiners. 1993. Pasture age effects on soil-atmosphere trace gas exchange in a deforested area of Costa Rica. Nature 365:244-246.

M. Keller, D.A. Clark, D.B. Clark, E. Veldkamp, and A. Weitz. 1996a. If a Tree Falls... Science (in press).

M. Keller, J. Melillo, and W.Z. de Mello. 1996b. Trace Gas Emissions from Ecosystems of the Amazon Basin. Ciencia e Cultura (in press).

Lal, R. 1981. Deforestation of tropical rainforest and hydrological problems. Pages 131-140 in Lal, R. and E. W. Russel (eds.), Tropical Agricultural Hydrology, J. Wiley, New York.

Lugo, A.E. and M.J. Sanchez. 1986. Land use and organic carbon content in some subtropical soils. Plant and Soil 96: 185-196.

Miller, R.L. and J.F. Cruise. 1995. Effects of suspended sediments on coral growth: evidence from remote sensing and hydrologic modeling. Remote Sensing of the Environment 53: 177-187.

Peterson, D.H. and J.F. Festa. 1984. Numerical simulation of phytoplankton productivity in partially mixed estuaries. Estuarine, Coastal, and Shelf Science 19: 563-589.

Rodriguez, M.I., M. Fries, L.Y. Garces, P. Steudler, J. Melillo, and G.A. Toranzos. 1993. Denitrifying bacteria

in temperate and tropical soils: occurrence of copper- and heme-type nitrate reductases. Proceedings of the 93rd General Meeting of the American Society of Microbiology, p. 74.

Salati, E., A. Dall'Olio, E. Matsui, and J.R. Gat. 1979. Recycling of water in the Amazon basin: an isotopic study. Water Resources Research 15: 1250-1258.

Sanchez, P.A. 1976. Properties and Management of Soils in the Tropics. J. Wiley, NY.

Shukla, J., C. Nobre, and P.J. Sellers. 1990. Amazon deforestation and climatic change. Science 247: 1322-1325.

Thomlinson, J. R., M.I. Serrano, T. del M. Lopez, T.M. Aide, and J.K. Zimmerman. 1996. Land-use dynamics in a post-agricultural Puerto Rican Landscape (1936-1988). Biotropica (in press).

Uhl, C. and C.F. Jordan, 1984. Succession and nutrient dynamics following forest cutting and burning in Amazonia. Ecology 65:1476-1490.

Uhl, C., R. Buschbacher, and E.A.S. Serrao. 1988. Abandoned pastures in eastern Amazonia 1. Patterns of plant succession. Journal of Ecology 76:663-681.

Uhl, C., D. Nepstad, R. Buschbacher, K. Clark, B. Kauffman, and S. Subier. 1990. Studies of ecosystem response to natural and anthropogenic disturbances provide guidelines for designing sustainable land-use systems in Amazonia. In "Alternatives to Deforestation: Steps Toward Sustainable Use of the Amazon Rain Forest" (A.B. Anderson, cd.), Columbia University Press, New York.

Wooldridge, R. 1986. Sedimentation in reservoirs: Magat reservoir, Cagayan valley, Luzon, Phillipines -1984 reservoir survey and data analysis. Report no. OD 69, Hydraulics Research Limited, Wallingford, UK, 67 pp.

Zou, X. and G. Gonzalez. 1996. Changes in earthworm density and community structure during secondary succession in abandoned tropical pastures, Soil Biology and Biochemistry (in press).

URC97115

Spectral Characterization and Geologic Mapping of the Middle Proterozoic Apache Group, Troy Quartzite, and Associated Diabase, Central Arizona, Utilizing Thermal Infrared Multispectral Scanner (TIMS) Imagery

115/43

JOHN M. SEELEY
Pan-American Center for Earth and Environmental Sciences
The University of Texas at El Paso, El Paso, Texas

Abstract

An integrated study employing a wide variety of geological, geophysical and remotely sensed data is underway to better understand the geologic history of Proterozoic rocks of central and southeastern Arizona. Middle Proterozoic-age reeks are well represented in this region by the Apache Group, Troy Quartzite, and coextensive diabase. This study utilizes thermal infrared multispectral scanner (TIMS) imagery for lithological mapping and spectral characterization of the Apache Group, Troy Quartzite and associated rocks. TIMS data is supplemented by Landsat Thematic Mapper (TM) data, SPOT panchromatic imagery, gravity and magnetic measurements, and conventional field mapping. Spectral image data will add mineralogic and lithologic information that cannot be obtained from the reflected image data alone. Results obtained will be combined to construct a concise, regional-scale photogeologic map of Proterozoic rocks of the region. Gravity and magnetic data will be employed for correlation studies, primarily between the Apache Group and the Unkar Group of the Grand Canyon.

Introduction and Regional Geology:

A major gap in our understanding of the tectonic evolution of North America is the Proterozoic history of the southwestern United States, The scattered outcrops of Proterozoic rocks in this region have eluded a convincing regional synthesis to date. Clearly, an integrated study employing a wide variety of geological, geophysical and remotely sensed data is required to better understand the geologic history of this region.

Middle Proterozoic-age rocks are well represented in central and southeastern Arizona by the Apache Group, the overlying Troy Quartzite, and coextensive intrusions of diabase (Shride, 1967, Wrucke, 1989). These rocks are well exposed between the Mogollon Rim and Globe, in the Mountain Region of central Arizona that belongs structurally to the Colorado Plateau Province (Shride, 1967). [n this region, the Apache Group and Troy Quartzite are broken by relatively few faults and lie in an approximately horizontal attitude. In the Basin and Range Province to the south, Proterozoic outcrops are scattered, structurally deformed and stratigraphically less complete. Rocks of the Apache Group and Troy Quartzite are found as far south as 32° latitude, from the Vekol Mountains to the west and the Little Dragoon Mountains to the east. The major outcrops of the Apache Group, Troy Quartzite, and the associated diabase intrusions are shown in Figure 1.

The Apache Group consists of, in ascending order, the Pioneer Shale, Dripping Springs Quartzite, Mescal Limestone, and an unnamed basalt. The total thickness of the Apache Group is about 490 m. The Pioneer Shale and Dripping Springs Quartzite are both comprised of basal conglomerates overlain by fine- to coarse-grained elastics with associated tuffaceous debris. Deposition is thought to have occurred in both fluvial and marine environments. The Mescal Limestone consists of lower and middle dolomitic members which are overlain by an argillite member probably tuffaceous in origin. Basalt flows are found locally below and above the argillite member. The Mescal Limestone attains a maximum thickness of approximately 130 m. Deposition occurred in shallow marine, intertidal and supratidal environments.

The Troy Quartzite consists of a lower arkose member, the middle Chediski Sandstone Member, and an upper quartzite member. The maximum thickness of the Troy is approximate y 365 m. The arkose was

deposited in fluvial and eolian environments, while the sandstones and quartzites are likely of marine origin, but interpretation remains problematic.

All of these rocks were intruded by diabase in sills ranging in thickness from a few meters to more than 400 m. The volume of diabase equals or exceeds the volume of strata in some areas. Due to sill emplacements, most of the dolomite of the Mescal Limestone was altered to silicated limestone, and some of the elastic rocks were converted to hornfels or quartzite.

Probable correlative rocks of the Apache Group and Troy Quartzite can be found in the Unkar Group of the Grand Canyon (Shride, 1967). Possible correlation between the Apache Group and the Pahrump Group of the Death Valley region may also exist (Wrucke, 1989). Rocks of similar ages and lithologies also exist in West Texas, however, correlation is speculative at best. Previous work (Seeley, 1990, 1996; Seeley, et. al., in review) regarding the stratigraphy and depositional environments of the Middle Proterozoic Lanoria Formation, located in the Franklin Mountains immediately north of El Paso, Texas, has provided the interest in continuing the study of depositional environments, correlative studies, regional tectonics, etc., of Middle Proterozoic rocks of the Southwest. Rocks of the Apache Group and Troy Quartzite provide an excellent opportunity for subsequent sedimentological studies involving depositional environments and paleogeographic reconstruction of the Southwestern United States during Middle Proterozoic time.

To date, there have been only a limited number of studies that have evaluated and utilized multispectral data for lithologic mapping (Gillespie, et. al., 1984; Abrahms, et. al., 1991; Hook, et. al., 1994; Crowley and Hook, 1996) and stratigraphic analyses (Lang, et. al., 1987; Ernst and Paylor, 1996). As thermal infrared data have been demonstrated effective for geologic interpretation, we plan to utilize thermal infrared multispectral scanner (TIMS) imagery for lithological mapping and spectral emissivity characterization of Middle Proterozoic rocks in central and southeastern Arizona. These data are supplemented by Landsat Thematic Mapper (TM) data, SPOT panchromatic imagery, gravity and magnetic measurements for regional correlation studies, and conventional field mapping techniques to provide additional geologic information.

Spectral Characterization

A major component of this study will contribute spectral reflectivity and emissivity data for the rocks of the Apache Group, Troy Quartzite and associated diabase. Spectral characterization of rock samples will utilize methods outlined by Lang, et.al. (1990). A flow chart for data analyses is shown in Figure 2. Following relevant field observation, image data review and literature review, a statistically valid sampling program will be developed. Representative rock and soil samples will then be collected from the study area. After collection, the samples will be split for archival storage and for actual analyses. Samples will be analyzed for their characteristic visible and infrared spectra at the Jet Propulsion Laboratory (JPL) in Pasadena, California. Field analyses may also be conducted with a portable spectrometer. Conventional X-ray diffraction analyses for mineralogical determinations will be conducted at the Department of Geological Sciences, University of Texas at El Paso. Infrared spectra will be sorted into spectral classes, based on the presence or absence of diagnostic visible and infrared reflectance and emission features.

Image Analyses and Interpretation

In the thermal infrared region (8-12 μm), spectral variations typically relate to differences in the Si-O bonding of silicate minerals. Multispectral thermal infrared images provide a way to distinguish rocks based on their silicate mineralogy, which is important in the classification of sedimentary and igneous rocks. The marked lithological contrasts between the various rocks of the region should provide for clear spectral differentiation for mapping purposes utilizing TIMS imagery.

The TIMS instrument provides 6-channel digital multispectral data in the 8-12 μm wavelength. Palluconi and Meeks (1985) describe the TIMS system in detail. The typical 5-km groundtrack and 25 m by 25 m resolution of TIMS data are considered adequate for making pictures suitable for photogeologic interpretation and mapping (Lang, et. al., 1987). Following acquisition, the TIMS data will be calibrated and corrected for atmospheric distortion by appropriate methods. Spectral classes will be displayed on false-color images for analyses.

Initial reconnaissance photogeologic interpretation will be conducted utilizing 1 :250,000-scale TM color-composite images and SPOT panchromatic images. These images will be used to identify regional stratigraphic markers and to locate undeformed reference localities for stratigraphic interpretation. Detailed photogeologic interpretation will be conducted at 1:24,000-scale for areas of interest. TIMS data will be digitally registered with the TM/SPOT data. Enhancement by color decor-relation contrast stretch (dstretch) methods (Kahle and Rowan, 1980, Gillespie, et. al., 1984), as initially described by Soha and Schwartz (1978) will be utilized to enhance spectral differences. Spectral data will add mineralogic and lithologic information that cannot be obtained from the TM data alone. Results obtained at 1:24,000-scale will be combined to construct a regional-scale photogeologic map. Preliminary results will be field checked as the mapping proceeds. It is hoped that the analysis of spectral imagery will provide additional information and improvement for previously mapped areas of the region.

We also plan to utilize the imagery to examine how lithologic changes (i.e. variations in quartz content), if apparent, within the Troy Quartzite relate to paleodispersal patterns of sediment. Sediment dispersal patterns, combined with geophysical studies, could be used to better understand the Middle Proterozoic tectonic setting of the region.

Gravity and Magnetic Data

Correlation of Proterozoic rocks of the southwest U.S. remains problematic. A major issue of this study is the possible correlation of outcrop of the Apache Group/Troy Quartzite found in scattered mountain ranges, with the exposures of equivalent age rocks found in the Grand Canyon. Gravity and magnetic data will be employed to detect structures in the Proterozoic rocks where they are covered by younger strata.

Summary

Results of this study will contribute significantly to our understanding of the Middle Proterozoic tectonic setting of the southwestern United States. The major objectives of this study are threefold: 1) to contribute spectral reflectivity and emissivity data for the rocks of the Apache Group, Troy Quartzite and associated diabase; 2) to provide a concise and accurate geologic map of the Apache Group/Troy Quartzite utilizing a variety of remotely sensed data; and, 3) to utilize gravity and magnetic data for correlation studies between the Apache Group/Troy Quartzite and Unkar Group rocks of the Grand Canyon.

References Cited

Abrahms, M. J., Abbott, E. A., and Kahle, A. B., 1991, Combined use of visible, reflected infrared and thermal infrared images for mapping Hawaiian lava flows: Journal of Geophysical Research., vol. 96, p. 475-484.

Crowley, J. K. and Hook, S. J., 1996, Mapping playa evaporite minerals and associated sediments in Death Valley, California, with multispectral thermal infrared images: Journal of Geophysical Research, vol. 101, no. B1, p. 643-660,

Ernst, W. G. and Paylor, E. D., 1996, Study of the Reed Dolomite aided by remotely sensed imagery, central White-Inyo Range, easternmost California: AAPG Bulletin, vol. 80, no. 7, p. 1008-1026.

Gillespie, A. R., Kahle, A. B., Palluconi, F. D., 1984, Mapping alluvial fans in Death Valley, California, using multichannel thermal infrared images: Geophys. Res. Lett., vol. 11, p. 1153-1156.

Hook, S. J., Karlstrom, K. E., Miller, C. F., and McCaffrey, K. J. W., 1994, Mapping the Piute Mountains, California, with thermal infrared multispectral scanner (TIMS) images: Journal of Geophysical Research, vol. 99, no. B8, p. 15605-15622.

Lang, H. R., Bartholomew, M. J., Grove, C. I., and Paylor, E. D., 1990, Spectral reflectance characterization (0.4 to 2.5 and 8.0 to 12.0 μm) of Phanerozoic strata, Wind River Basin and southern Bighorn Basin areas, Wyoming: Journal of Sedimentary Petrology, vol. 60, no. 4, p. 504-524

Lang, H. R, Adams, S. L., Conel, J. E., McGuffie, B. A., Paylor, E. D., and Walker, R. E., 1987, Multispectral remote sensing as stratigraphic and structural tool, Wind River Basin and Big Horn Basin areas, Wyoming, AAPG Bulletin, vol. 71, p. 389-403.

Palluconi, F. D. and Meeks, G. R., 1985, Thermal infrared multispectral scanner (TIMS): an investigators guide to TIMS data: Jet Propulsion Laboratory Publication 85-66, 68p.

Seeley, J. M., 1990, Stratigraphy, petrology, and depositional environments of the Middle Proterozoic Lanoria Formation, Franklin Mountains, El Paso County, Texas: [M.S. Thesis], Texas A&I University, Kingsville, Texas, 134 p.

Seeley, J. M., 1996, Depositional environments of the middle Proterozoic Lanoria Formation, Franklin Mountains, West Texas: (abs.), Southwest Section American Association of Petroleum Geologists Convention, El Paso, Texas.

Seeley, J. M., Thomann, W. F., and Marsaglia, K. M., (in review), Sedimentation and associated depositional environments of the Middle Proterozoic Lanoria Formation, Franklin Mountains, West Texas: Sedimentary Geology.

Shride, A. J., 1967, Younger Precambrian Geology in Southern Arizona: Geological Survey Professional Paper 566, 89p.

Wrucke, C. T., 1989, The Middle Proterozoic Apache Group, Troy Quartzite and associated diabase of Arizona, *in*, Jenney, J. P. and Reynolds, S. J., eds., Geologic Evolution of Arizona, Arizona Geological Digest 17, p. 239-258.

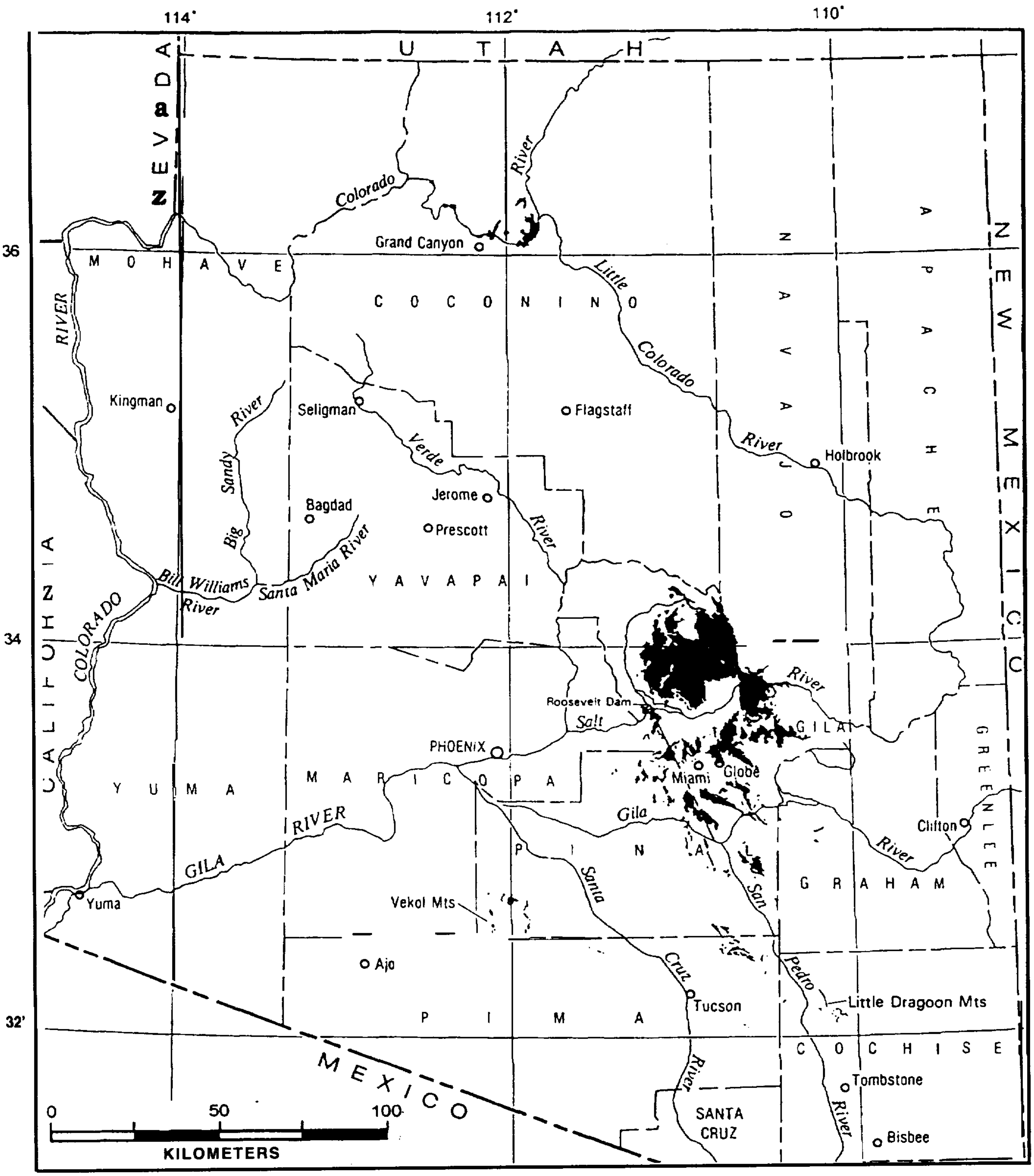


Figure 1: Regional map showing outcrop areas of the Apache Group, Troy Quartzite and associated diabase in southern and central Arizona. The Unkar Group in the Grand Canyon, located in northern Coconino County, is also shown. Map modified from Wrucke, 1989.

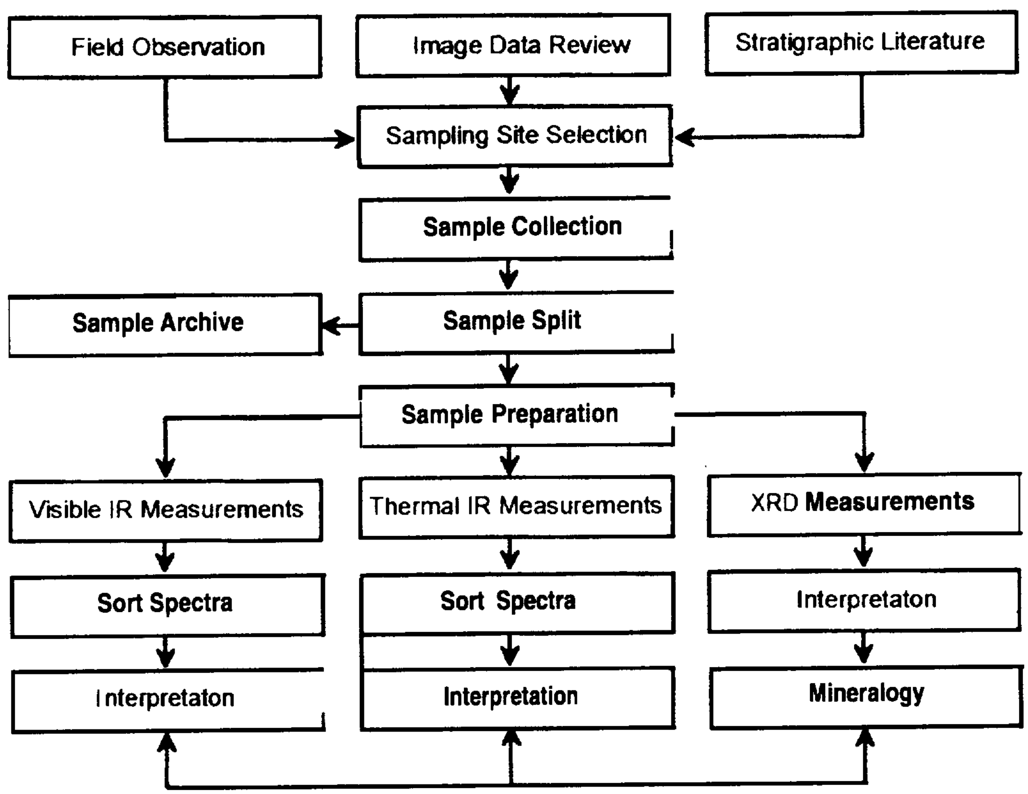


Figure 2: Flow chart illustrating proposed approach to be used to obtain spectral characteristics from Proterozoic rocks in the study area (modified from Lang, et. al., 1990).

N A S A 5116 / 30

URC97116

Contributions to Educational Structures that Promote Undergraduate Research

JOHN SEPIKAS, MILAN MIJIC,
DON YOUNG and STEVE GILLAM

Abstract. The opportunities for community college and traditionally underrepresented minority students to participate in research experiences are typically rare. Further, what research experiences that are available often underutilizes the students' potential and do not have follow-up programs. The Physics Outreach Program (POP) working in conjunction with the Jet Propulsion Laboratory is designed to reach out to this segment of the student population and encourage them to consider careers in physics and astronomy. The program is special in that it creates a "vertical" consortium or pipeline of schools whereby students graduating from one participating institution will then transfer to another. This helps to insure that participating students will experience continuity and, with the assistance of JPL equipment and staff, a quality of instruction that they would otherwise not be able to afford.

Key words. educational outreach, undergraduate research, community college research, underrepresented minority student research

I. Program Overview

The Physics Outreach Program (POP) is a consortium of local Los Angeles universities, community colleges, high schools and NASA's Jet Propulsion Laboratory (JPL) [1]. Its primary purpose is to encourage students, especially underrepresented minorities, to consider physics and astronomy as career choices. A secondary purpose is the establishment of a direct mechanism for transfer of research methodology from JPL to the local participating science education departments and the strengthening of the cooperative ties between those departments. Last] y,

the formal results of the research have direct application to the mission design of current or proposed flight projects at JPL.

To achieve its primary purpose the POP program has adopted a "pipeline" or vertical organizational structure. The idea here is that students would go to, and be supported from, high school to community college, to state college and then finally to graduate school. If all goes well some of those students could end up as contributing professionals at research institutions.

Communication among faculty along such a structure can strongly promote improvements in curricula, educational delivery and laboratory experiences at the entry level institutions. Faculty at such institutions can then more strongly justify requests for course and laboratory upgrades with school administrators with the argument that they are required to maintain the level of quality expected of the transferring institution and JPL. This type of structure could serve as one possible model to improve science education nationally in conjunction with NASA field centers and their affiliated local universities.

Under ideal circumstances potential physics and astronomy students would enter the program at participating high schools, typically at the beginning of their junior year. For their participation, they would receive financial support at a level consistent with entry level high school employment. This would in turn obligate them to stay in school, maintain a B average in their science classes, participate with the college students in field training exercises and generally expose them to the potential of the program. The financial support is crucial to establish the commitment of the program *for* the student and to be able to demand a higher level of responsibility *from* the student that would in general not be required in the high school environment.

The program is committed to communicate to the student that they have stable and positive choices in life, and that they can join a nurturing community and environment that understands and appreciates their desire to do science. Such an environment is important to all young

science students but especially to minority students who traditionally lack exposure to a science community.

At the end of their high school participation students should have established their desire to a career in the sciences or that science is a direction they wish to pursue. And although their choice may not necessarily be in physics and astronomy that they are planning to enroll in the standard science/engineering courses. Additionally, they should have received enough instruction and training with JPL procedures, protocol and equipment so as to be productive when they enter into the next phase in the collegiate program.

From high school, again under ideal circumstances, students would transfer to a local community college or in special cases directly to a participating state university. The community college transfer choice has the advantage of offering the student a more gradual incorporation into collegiate life and a tremendous cost savings. Students would receive the same stipend that they would have received at the state university so as not to discourage them from attending a community college solely on the basis of discrepancies in financial support. Such savings could potentially be used for graduate school. Additionally, the community college is sufficiently close in most cases so as to allow the student to remain living at home so as to achieve further saving and allow more time for emotional growth. Further, academic deficiencies that may be present can also be addressed.

At the end of the community college experience the student should: have all lower division science requirements completed, have sufficient experience with JPL equipment to have been checked out on all telescopes and to conduct observational investigations without direct supervision, and have declared as their major physics or astronomy.

From the community college, yet again under ideal circumstances, transfer to a participating state college would occur. Students would complete their baccalaureate degrees in physics or astronomy, write a paper to be published as their senior thesis and apply to graduate school.

II. Program Experience

As mentioned before the above overview represents the ideal case. The experiences of the past two summers are distilled in this section.

The previous overview assumes that the program was year-round. Unfortunately funding was available only for the summer research component for the summers of 1995 and 1996. The participating schools were Pasadena Unified School District GeoSpace Academy (sponsored by JPL and known simply as Space Academy), Pasadena City College (PCC), Los Angeles City College (LACC) and California State University, Los Angeles (CSLA). All of the above schools are minority serving institutions.

Since part of the funding for the program came from the Galileo, Cassini and the Mars Pathfinder flight projects offices respectively, research direction came from scientists on these projects. All work was done at JPL's Table Mountain Observatory (TMO) which is about a two hour drive (in good traffic conditions!) from downtown Los Angeles. The telescopes that were used are both of Cassegrain configuration with 1.2 meters and .6 meters of aperture. CCD's used for the observations were both thermoelectrically and LN2 cooled. Data reductions were done on Sun work stations using IRAF (Image Reduction and Analysis Facility) or on 486 based PC's using generic image processing software.

The summer of 1995 was the pilot year of the program, with just PCC and CSLA conducting active research. The following is a brief summary of those efforts.

The PCC team concentrated on faculty/infrastructure development and observationally, at the request of the Cassini project [2], timed the occultations between the moons of Mimas and Enceladus. Two methodologies were employed. One was to construct the combined light curve of both moons and deduce the time when the curve first dropped as the time of first contact. The other was to plot the separation between

the moons as a function of time and then to extrapolate to the time when that separation was zero as the time of maximum occultation. 1995 was a rare triple ring plane crossing for Saturn and hence an especially good time to do occultation timings. These observations were part of an international effort to do such timings and the data was used to upgrade the ephemerides of these moons to improve the navigational capabilities of the Cassini spacecraft when it orbits Saturn.

The CSLA group worked on three projects [3], [4],[5]: general scans of the atmosphere of Jupiter to look for remnants of the Shoemaker-Levy 9 impact to support observations of Galileo in orbit about Jupiter, data reductions of previous observations of Mars to characterize the impact of dust storms on potential landing sites, creation of an educational movie which shows the motions of the moons of Saturn, and integration of an infrared spectrometer on the 1.2 meter telescope.

In 1996, during the bad weather season, work started on two instrumentation projects: installation and testing of the conditioners for decreasing instrumental noise on the 1.2 meter telescope CCD camera, and installation and testing of LN2 cooled near infrared imaging spectrometer. The later project also included design of an optical and mechanical system for the simultaneous imaging in the visual and near IR.

For the summer of 1996 students from the Space Academy and LACC joined the program. The research directions were observations of comet Hale-Bopp, the rings and atmosphere of Saturn, continued development of the infrared spectrometer and astrometry of Jupiter's outer moons.

III. Future Developments

Additional community colleges are planned to join the program in 1997. They are East Los Angeles Community College (ELAC) and South West Community College (SWCC). Both are minority serving institutions which have a cluster of minority high schools sending students to them. Since not all students will elect(unfortunately!) to continue in physics and astronomy these additional schools will assist in recruitment and retention efforts, The hope is to create a critical mass sufficient to have students ultimately complete their baccalaureate at CSLA and gain the recognition of prospective Ph.D. granting institutions as a place for them to find well qualified students with research experience.

In the Pasadena area, the Pasadena Unified School District along with the California Institute of Technology has implemented a "pre-academy" at the junior high level as a primer for the existing space academy. This also complements the Caltech SEED program to develop science education at the elementary school level. Thus through public education in Pasadena, students can in the sciences, go from grammar school to graduate school. This again illustrates the advantage of the "pipeline" structure.

The majority of the funding for the current program was provided by NASA's office of Equal Opportunity, Code E. The success of the program can hopefully be continued to a permanent year round program through finding from the National Science Foundation modeled after the successful program currently in place at the University of Wyoming. Continuity and consistency are key for the POP program, or any program like it, to achieve the goal of increasing the number of students, especially underrepresented minority students, receiving Ph.D.'s in physics and astronomy.

The authors would like to thank NASA and the various JPL project and program offices for their support of this program.

References:

[1] D. Young, et al, *Los Angeles Minority Science Consortium,* Internal NASA proposal, February, 1995.

[2] J. Sepikas, C. Gillingham, G. Miller, and Andy Hammond, *Final report to Project Cassini,* November, 1995.

[3] J. Atienza-Rosel and D. Moreno, *"Jupiter in August* 1995", CSLA report to the Galileo mission, November, 1995,

[4] H.A. Avila, C. Russel and J. Atienza-Rosel, *"Mars on February 17, 2995",* CLSA report to Mars Pathfinder mission.

[5] H.A. Avila and D. Moreno, *"Dust storms on Mars",* poster paper on CSLA conference on undergraduate research, 1996.

Page intentionally left blank

/// /' ~ ~//7/63

URC97117

Transformed Vector Quantization Using A Neural Network Approach

E. Sherrod and Robert Li
NASA ACE Center
Department of Electrical Engineering
N.C. A&T State University
Greensboro, N.C. 27411

N. Al-shamakhi
Vocational Training Institute
Ibri, Oman

Keywords: Image compression, Neural network, Vector quantization, Frequency transform

Abstract

Digital image compression is an important technique in digital image processing. We have used a modified competitive learning algorithm to achieve big saving in training time. Moreover, we investigate a TVQ (transformed vector quantization) method to achieve the highest possible compression ratio without much sacrifice in image quality.

1.0 Introduction

The image compression techniques can be used in areas of high information volume to reduce the data rate to within the channel capacity. Large amounts of compression can be obtained with lossy compression techniques. One of the most important lossy compression approaches is the VQ (vector quantization) method[1,2,3]. The VQ approach is basically a clustering method. The vectors are obtained from image data by extracting nonoverlapping blocks and arranging the pixels in each block in a line by line order. The VQ method builds up a codebook, a dictionary of a few representative vectors (codevectors). Then each block in the image is coded with the index value of the closest codevector in the codebook. The signal-noise-ratio (SNR) is usually used to measure the fidelity or quality of recovered image, By using SNR as a quality criterion, one can compare the performance of different coding schemes with the same codebook size, An important advantage of VQ image compression is its fast decompression by table lookup technique. In our work, we are using an improved version of competitive learning neural network to speed up the codebook design process. More importantly, we are developing a technique called TVQ (transformed vector quantization) to achieve very high compression ratio while maintaining a reasonable image quality.

2.0 Competitive Learning Algorithm for Vector Quantization

Neural network approaches have been used for pattern classification and data clustering [4,5]. It is possible to apply the training algorithm of neural networks to the design of appropriate codebook which maximizes the SNR values of reconstructed image. One approach is based on competitive learning algorithms. The basic idea is that only one

output unit of the neural network is active(on) at a given time, The output units compete for the active status, and are often called winner-take-all units. The model can be used to categorize or cluster the input data [9], This leads to its usefulness in image vector quantization. I n our implementation, a given image is first segmented into vectors of the same size. The elements of those vectors are gray levels of the pixels in the image. A neural network is set up with an input and output layer. For the input layer, its number of input nodes equals to the dimensions of image vectors. For the output layer, its number of output nodes equals the size of the codebook. Recently, we have introduced a modified competitive learning algorithm to improve the speed of codebook design, to resolve the problem of dead units, and to obtain a superior image quality for decompressed images[10], However, these improvements are achieved in time-domain. In this project, we are investigating the possibility that further improvement can be obtained by vector quantization in the frequency domain.

3.0 Transformed Vector Quantization(TVQ)

Prior to forming the vectors of an input image for VQ, the original image data can be mapped to the frequency domain by an application of a transform such as DCT. The produced transformed coefficients are then used as vector components and can be vector quantized. After vector quantizing, the quantizer output are inverse transformed to produce the quantized approximation of the original input image [7.8], This idea is illustrated in Figure 1, where X represents the original vector, Y is the transformed vector. Y is the quantized output, and $\dot{X}$ is the quantized approximation of X.

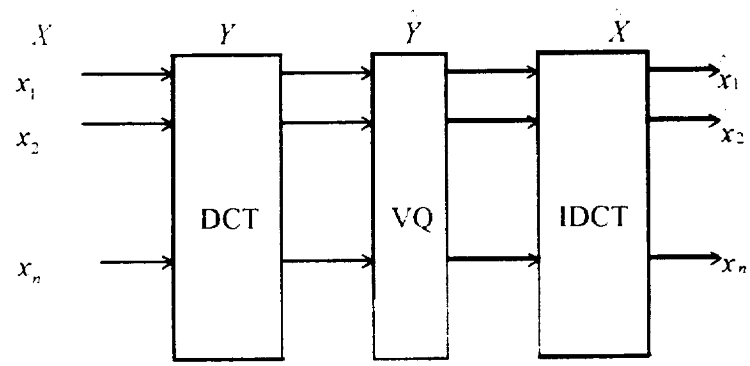


Figure 1 A transformed VQ system

One might ask what is the advantage of combining transformation and VQ? In fact, if the dimension of TVQ is of the same size as that of VQ, there would be no advantage gained from the transformation. This is because the optimal code vectors for the combined method (TVQ) would be exactly same as that of direct VQ. The average distortion error in both cases would be identical [1]. However, there is an advantage that can be obtained by using TVQ. To understand how this is possible, we need to consider the transform coefficients and their distribution in the frequency domain. When a linear transform is applied to the original vector signal, the information is compacted into a

subset of the vector components. DCT maps data from the spatial domain to the frequency domain which often results in that the high energy components would be concentrated in the low frequency region, This means that the transformed vector components in the higher frequency regions have very little information. These low energy components might be treated in the following way: Discard these components entirely and hence the dimension of transformed vectors and the complexity of VQ are both reduced, This yields a reduced codebook size which means a higher compression ratio than that of VQ alone. Figure 2 shows a truncated TVQ system where the truncation El, takes a k-dimensional input vector, Y and maps it into a p-dimensional ($p < k$) vector, Z. Then p-dimensional VQ is used to vector quantize the truncated vector, forming the quantized approximation $\hat{Z}$ Recovery can be obtained by a padding operation, H*, that appends (k-p) additional components with value zero to the vector, producing the k-dimensional vector Y. Finally, inverse transform is applied to $\hat{Y}$ to obtain the approximation $\hat{X}$.

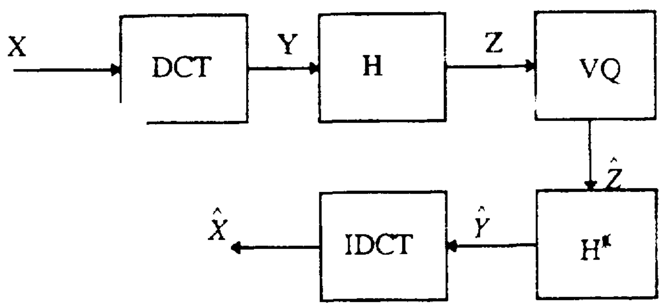


Figure 2: A truncated TVQ system

In our work, after transforming the image vectors to frequency domain using DCT, the low energy components which count a certain percentage of all transformed components, were truncated from each subimage leaving only the high energy components. The truncated transformed image was then vector quantized. The following is a list of steps for the TVQ algorithm:

<u>Step 1:</u> The image was transformed to frequency domain using DCT.
<u>Step 2:</u> The low energy components were truncated from each subimage.
<u>Step 3:</u> Using the truncated transformed image, the modified competitive learning algorithm was performed for vector quantization using subimages of size n by n.
<u>Step 4:</u> At the receiving end, after decoding the codewords, the data was inversely transformed using IDCT to obtain an approximated image.

The complexity saving with this truncation technique can be quite substantial. To understand this, suppose that a resolution of r bits per vector component is allocated to code the k-dimensional vector X for a total bit allocation of rk bits. Without truncation the maximum possible codebook size would be $N = 2^{rk}$ With truncation and using the same resolution r, the maximum possible codebook size is reduced to $N = 2^{rp}$, which

can be orders of magnitude smaller depending on the values of k and p. When k is large, the savings in complexity can be very substantial while degradation (due to truncated energy) can be negligible. In other words, TVQ with truncation can become a good complexity reducing technique for image coding. Alternatively, for the same complexity, the resolution or codebook size can be increased to obtain better performance.

4.0 Experimental results

TVQ was done using DCT with a block size of 8 x 8. Table 1 shows the results for the "bridge" image using modified VQ and TVQ. Modified VQ refers to the VQ method using the modified competitive learning algorithm reported in [10].

Table I

VQ and TVQ results for the "bridge" image

Algorithm	codebook size	Bit rate (bpp)	CR	SNR (dB)	CPU time (See)
VQ	50	0.088	90.7	28.2	24
TVQ	50	0.088	90.7	30,4	9.1
VQ	25	0.073	110.3	27.1	6,4
TVQ	25	0,073	110.3	29.6	4.7
VQ	15	0.061	131.1	26.7	2.9
TVQ	15	0.061	131.1	28.5	2.6

From Table 1, we can see that both methods have produced higher compression ratios while the quality of recovered images has been maintained within an acceptable level. Using a block size of 8 x 8 and the same codebook size (at 50, 20 and 15), TVQ proved to be better than VQ in terms of image quality and training time. The reason for this is that the truncated transformed image has less complexity in sample space. Therefore, fewer output nodes are needed to recover the image than that required by VQ. Figures 3 and 4 show the recovered images obtained by the VQ and TVQ algorithms, respectively, It can be seen that TVQ result emphasized low-frequency information while the overall SNR ratio is better than that of VQ.

5.0 Summary

The goal of data compression is to reduce the bit rate for data storage or transmission while maintaining an acceptable image quality. This research investigates a new approach to vector quantization technique for image compression TVQ approach starts by first

transforming image data to the frequency domain using discrete cosine transform. Then within a transformed sub image, the high-frequency components were truncated. As a result, the truncated subimage was vector quantized using a smaller dimension for its vector, thus reducing the complexity of the sample space. Based on the results obtained thus far, TVQ can achieve very high compression ratio without much sacrifice in image quality.

Acknowledgments

This research was partially supported by the NASA Autonomous Control Engineering (ACE) Center under grant ACE-48146.

Reference

1 A. Gersho and R, Gray, *Vector Quantization and Signal Compression,* Kluwer Academic Publisher, 1992.

2. A. K. Jain,"Image data compression: a review, " *Proc. IEEE,* vol. 69, pp. 349-389.

3. A. N. Netravali and J. O. Limb, "Picture coding: a review," *Proc. IEEE*, Vol 68, No. 3, pp. 366-406.

4. R. P. Lippman, "An introduction to computing with neural nets", *IEEE ASSP Magazine,* April 1987, pp.4-21.

5. D. E. Rumelhart, et. al., *Parallel Distributed Processing*, The MIT Press (1 986), vol. 1.

6. J. Hertz, A. Krogh, and R. Palmer, *Introduction to the theory of neural computation,* Vol. I of Santa Fe Institute Lecture Notes, Addison-Wesley Publishing Co., 1991.

7. "R. Gonzalez and R. Woods, *Digital Image Processing,* Addison-Wesley, 1992

8. Y. Linde, A. Buzo and R. M. Gray, " An algorithm for vector quantization design, " *IEEE Trans. Communication,* Vol. com-28, pp. 84-95, Jan, 1980.

9. T. Kohonen, *Self-organization and associative memory,* 3rd Edition, New York, Berlin: Springer-Verlag, 1989,

10. R. Y. Li, E. Sherrod, G. Pan, and J. Kim, " Fast Image Vector Quantization Using A Modified Competitive Learning Neural Network Approach', to appear in *Int. Journal Of Imaging System and Technology.*

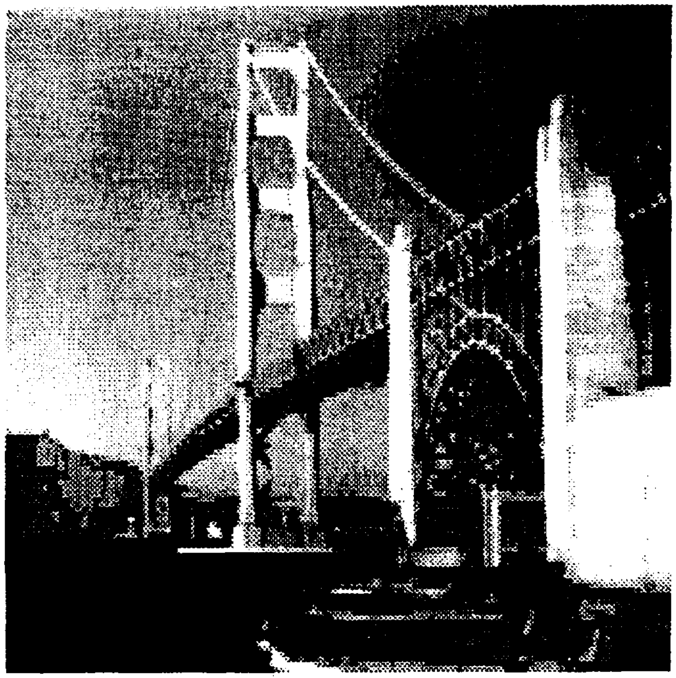


Figure 3: VQ “bridge” result (codebook size = 50)

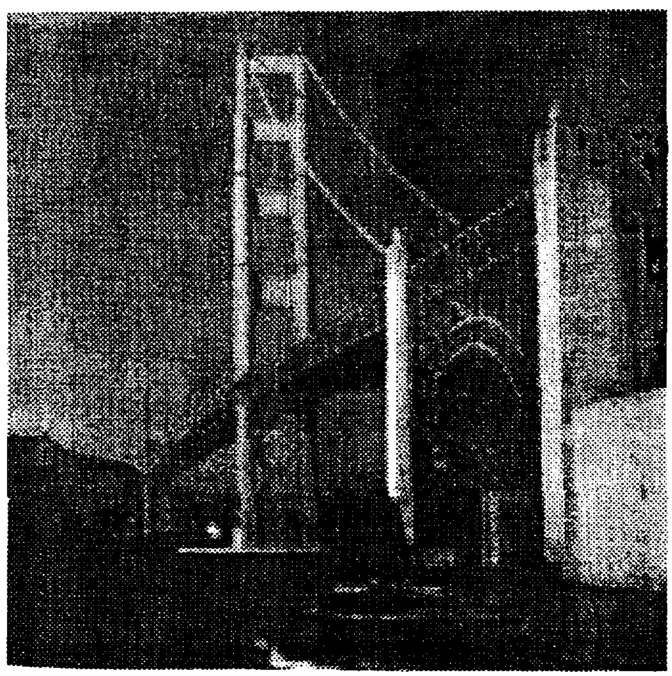


Figure 4 TVQ “bridge” result (codebook size = 50)

N11-87 5113/80

URC97118

The PACES Summer Science Trek; A Pre-College Science

Michelle B. Smith

Pan American Center for Earth and Environmental Studies (PACES)
University of Texas at El Paso
El Paso, Texas 79968
e-mail: michsmi@mail.utep.edu

Introduction

The University of Texas at El Paso (UTEP) received five-year funding to form the Pan American Center for Earth and Environmental Studies (PACES) in July 1995. PACES has as its goals to conduct research contributing to NASA's Mission to Planet Earth and to develop skilled scientists and engineers. PACES seeks to gain a more comprehensive understanding of geological, ecological and environmental processes and changes taking place in the southwestern United States and northern Mexico region. The PACES center has collaborative ties with two NASA field centers (Goddard Space Flight Center and Ames Research Center) and the Jet Propulsion Laboratory.

The original proposal contained no provision for outreach programs. However, at a meeting in the fall of 1995, Dan Goldin, NASA Administrator, issued the challenge that in order to accomplish NASA's goals to educate more of the citizenry in science and engineering, the Centers should take a broader perspective aimed at younger children.

Objectives

The objective of our program was to develop and support an interest by young women in science, mathematics, and pre-engineering coursework and to increase student awareness of the academic preparation necessary for such careers. Enrichment activities in the physical sciences, mathematics, pre-engineering and computer science, involved the use of simple equipment, toys, modern teaching techniques and the latest educational technlogy to stimulate the students. Additionally, the students were introduced to the environment and resources of the college and research organizations.

Participant Recruitment and Selection

Considering that women and minorities are emerging as significant talent pools in science, mathematics and engineering, and that they continue to be underrepresented in their fields, our program was aimed at young girls enrolled in Grades Six through Eight. The Girl Scouts were selected as our first group for two reasons. Dr. France Cordova, former Chief Scientist of NASA, once made the statement that of all the women she'd known at NASA, over 75% were former Girl Scouts. It was also suggested that a one-gender group might be easier to manage for our first effort. The Center coordinated its efforts with the Director of Program Services, Ms. Janet Brown, at the Rio Grande Girl Scout Council.

Initial participation in the summer outreach program involved about 30 Girl Scouts. Participants were held to an even number as the girls were partnered to share equipment and experiments. To screen the applicants, a form was developed which would allow candidates to be selected that perhaps had an undeveloped aptitude for science and would also enable statistical information requested by NASA to be collected.

Summer Program

The core of the one-week summer program was a laboratory oriented hands-on experience in the physical sciences, engineering and computer science. The students were exposed to optics, mechanics, and electricity through the use of very simple equipment and toys (winding toys, balloons, roller skates, etc.) The total cost for the supplies for the program was approximately $1,000. One-half of the expense was for rockets and engines which were built and launched on the last day of the program. Although new rockets will have to be ordered for the next group, the toys will be used for years to come, thus reducing the cost for subsequent years. An assortment of common household items such as straws, beads, straight pins and glue were also used. Using the toys grabbed the attention of the students, created the need to learn something new, provided the summary for closing a topic or unit, and facilitated easy integration into the curriculum.

Physical Facilities

The University of Texas at El Paso provided the project with excellent modern physical facilities. The Computer Science Department donated the use of the Computer Lab, providing individual state of the art workstations for each participant. The participants genuinely enjoyed exploring the Internet with minimal supervision. Each individual was allowed to handle rocks and minerals from the Geology Department specimen lab, and the Engineering Department donated supplies and materials for the electrical experiments. Our participants were given guided tours of the various colleges and community facilities on campus during the lunch break.

Project Staff

Dr. Andres Rodriguez, Professor of Physics of the University of the Pacific in Stockton, California. was engaged as a consultant to assist with the development of the curriculum and to conduct the first program. Dr. Rodriguez is widely recognized as a leader in minority outreach programs and has received numerous grants for such efforts from the U.S. Department of Education and the Department of Energy. He is also a consultant for the Exploratorium Science Museum in San Francisco. Dr. Rodriguez was assisted by the PACES Program Coordinator, Michelle Smith. Three senior electrical engineering students were recruited to provide daily instruction and personal attention. Additional UTEP students were recruited for selected demonstrations which included robotics, geology, and computer technology. Their presentations, together with interactions with our participants, were intended to motivate the children to learn more about careers and responsibilities for the future. The presence of successful undergraduate students as teaching assistants and mentors helped encourage the participants to think about their own college education at an early stage in their lives.

Summary and Future Work

The young girls who participated in our program gained knowledge and confidence in science and how it relates to their every day lives. The students were exposed to many different aspects of the science and engineering fields in order to encourage them to explore these fields for possible future careers. Experience gained from our initial program will enable the Center to fine-tune the program to provide an intesting and exciting summer enrichment program in the years to come.

Acknowledgments

We wish to thank NASA for its support of this project through the grant NCCW-0089. We also wish to extend our gratitude to Dr. Andres Rodriguez of the University of the Pacific for his guidance and kind understanding, and Janet Brown, for getting our program off the ground with the Girl Scouts.

/// P 319/32

URC97119

LIGHT-INDUCED ALTERATIONS IN STRIATAL NEUROCHEMICAL PROFILES

Angela E. Sroufe, J.A. Whittaker, and J. W. Patrickson.
Neuroscience Institute and Department of Anatomy
Morehouse School of Medicine
Atlanta, GA 30310

Introduction

Much of our present knowledge regarding circadian rhythms and biological activity during space flight has been derived from those missions orbiting the Earth. During space missions, astronauts can become exposed to light/dark cycles that vary considerably from those that entrain the mammalian biological timing system to the 24-hour cycle found on Earth. As a spacecraft orbits the Earth, the duration of the light/dark period experienced becomes a function of the time it takes to circumnavigate the planet which in turn depends upon the altitude of the craft. Orbiting the Earth at an altitude of 200-800 km provides a light/dark cycle lasting between 80 and 140 minutes, whereas a voyage to the moon or even another planet would provide a light condition of constant light (Stampi, 1994). Currently, little is known regarding the effects of altered light/dark cycles on neurochemical levels within the central nervous system (CNS).

Many biochemical, physiological and behavioral phenomena are under circadian control, governed primarily by the hypothalamic suprachiasmatic nucleus. As such, these phenomena are subject to influence by the environmental light/dark cycle. Circadian variations in locomotor and behavioral activities have been correlated to both the environmental light/dark cycle and to dopamine (DA) levels within the CNS. It has been postulated by Martin-Iverson et al. (199 1 & 1992) that DA's role in the control of motor activity is subject to modulation by circadian rhythms (CR), environmental lighting and excitatory amino acids (EAAs). In addition, DA and EAA receptor regulated pathways arc involved in both the photic entrainment of CR and the control of motor activity. The cellular mechanisms by which DA and EAA-receptor ligands execute these functions, is still unclear. In order to help elucidate these mechanisms, we set out to determine the effects of altered environmental light/dark cycles on CNS neurotransmitter levels. In this study, we focused on the striatum, a region of the brain that receives a number of dopaminergic and glutamatergic input and is known to be involved in the modulation of locomotor and behavioral responses.

Methods

Animal Entrainment and Tissue Extraction. Adult male Long-Evans Hooded rats (200-250 g) were maintained on a 12 hour light/dark (12L:12D) cycle (lights on at 0600 hours; lights off at 1800 hours) in laboratory quarters with free access to food and water. After 3 days of synchronization, rats were exposed to constant dark or constant light conditions for 24 and 48 hours, respectively. Control animals remained in the 12L:12D. After 24 and 48 hours, animals were decapitated either one hour prior to the onset of day light (subjective night) or one hour prior to the onset of night (subjective day). These times to sacrifice the animals were chosen to allow for the possible accumulation of the metabolic changes over the light or dark period. The brains were rapidly removed, submerged in ice-cold 0.9% saline and dissected on ice. Wet tissue weights were determined and sections were frozen at -80°C. Frozen sample aliquots from striatum were sonicated on ice in 0.1 M $HClO_4$ (1 0μL/mg wet tissue weight) for three 10-second intervals. The homogenate was centrifuged at 12,000 rpm for 30 minutes at 4°C.

Dopamine and Glutamate Analysis. Monoamine and amino acid levels were determined simultaneously according to the method of Gamache et al., (1993). Analytes were separated on an HR-80 column (ESA, Inc., Bedford, MA) and maintained at 33°C with mobile phases flowing at 1.0 mL/min. Four serial coulometric electrodes with applied potentials of 70, 140, 210 and 280 mV were used for the measurement of dopamine, while four serial electrodes all set at 450 mV were used for analysis of o-phthalaldehyde-β-mercaptoethanol (OPA-βME) derived glutamate. For dopamine analysis, 20 μL of the supernatant was injected into a high-performance liquid chromatography (HPLC) system with 8 electrochemical sensors (CEAS model

5500, ESA, Inc., Bedford, MA). The monoamine mobile phase consisted of 0.05 M monobasic sodium phosphate; 250 mg/L heptanesulfonic acid; 8% methanol v/v and was adjusted to pH 3.0 with HPLC grade phosphoric acid. For glutamate analysis, 30 µL of supernatant was combined with 50 µL of OPA-βME derivatizing solution and 20 µL was injected into the HPLC for analysis. The amino acid mobile phase consisted of 0.1 M dibasic sodium phosphate, 3.1% actonitrile (v/v) and 25% methanol (v/v) adjusted to pH 6.8 with phosphoric acid. OPA stock was prepared weekly. 27 mg of OPA was dissolved in 1.0 mL of methanol. Following the addition of 5 µL of βME, the solution was diluted to 10 mL with a borate-EDTA solution containing 0.1 M sodium tetraborate (pH 9.3) and 10 µM EDTA. The stock solution was stored at 4°C and protected from light. The working solution was prepared daily by diluting the stock solution with 3 parts borate/EDTA solution and placed into the refrigerated autosampler in an amber vial.

Kynurenic Acid (KYNA) Analysis. 20 uL of the supernatant was injected into the HPLC for analysis. KYNA levels were measured according to an isocratic method. The mobile phase consisted of 50 mM Na_2HPO_4 in 5% methanol at pH 6.26 (adjusted with H_3PO_4). A 15 cm x 4.6 mm^3 um C_{18} reverse-phase column (HR- 150, ESA, Inc.) with a flow rate of 1 mL/min was used. The electrochemical sensors were set at 150,200,300,600, 750, 1000, 1040 and 1090 mV.

Results

Striatal DA levels following continuous light exposure are shown in Figure 1. DA levels during the subjective night (samples taken at 5:00 a.m., est) were elevated relative to controls after 24 hours. However, following 48 hours of continuous light DA levels were suppressed. When compared to 12L:12D controls, animals held in 24 or 48 hours of continuous light registered no change in DA levels during the subjective day (samples taken at 5:00 p.m., est). Striatal dopamine levels following continuous dark exposure are shown in Figure 2. In 24 hours of constant dark, an increase in DA was observed in both daytime and nighttime levels. This increase was more pronounced in the night-time levels. In both subjective day and subjective night, DA levels remained elevated at 48 hours of continuous dark.

Striatal glutamate levels following continuous light exposure are shown in Figure 3. During both the subjective day (5:00 p.m.) and subjective night (5:00 a.m.), glutamate levels peaked at 24 hours of exposure and then decreased at 48 hours. A similar profile was observed during constant dark shown in Figure 4. Clearly, there was a faster rise to peak levels during both the subjective day and subjective night at 24 hours as compared to constant light.

Since glutamate levels were found to vary, attempts were made to determine whether similar changes were occuring in other endogenous EAA receptor ligands. Thus, kynurenic acid levels in rat striatum were measured during constant light and the results are shown in Figure 5. Exposure to constant light conditions suppressed KYNA levels during both the subjective day and subjective night. However, KYNA levels did not vary significantly during the subjective night. During the subjective day, KYNA levels decreased significantly within the first 24 hours. These light-induced levels remained low as long as continuous light conditions were maintained. In contrast, an increase in striatal KYNA levels was observed within 24 hours of constant dark conditions (Figure 6). This increase was transient however, in that there is a decrease to initial control values within 48 hours. Subjective night-time levels did not vary in the first 24 hours of constant dark, however, these levels rose rapidly to reflect daytime levels within 48 hours. Clearly, at the end of 48 hours, striatal KYNA levels were low during constant light and are elevated to approximately day-time control levels in constant dark.

Discussion

In the present study, we attempted to characterize the alterations in neurotransmitter levels within the striatum associated with variations in the environmental light/dark cycle. In 1993, Engber et al. suggested key roles for DA and EAA receptor-regulated pathways in the control of motor responses. They suggested that DA receptor-mediated pathways exhibit varying degrees of sensitivity to EAA receptor blockade and that some DA receptor-mediated responses require simultaneous NMDA receptor stimulation. These findings agreed with earlier reports by Martin-Iverson et al. (1991& 1992) which suggested DA and EAA involvement in both locomotor

activity and the photic entrainment of circadian rhythms. However, the mechanisms by which dopaminergic neurons repond to alterations in the light/dark cycle is currently not known.

Results from these experiments contribute to a number of implications with regards to the response of the striatum to alterations in environmental lighting. Clearly, striatal DA, glutamate and KYNA undergo alterations in response to constant lighting conditions. Although a number of reports exist regarding glutamate's role in excitotoxic damage within the basal ganglia, little data is available explaining its impact on other transmitters with the region. Based on experimental evidence from Difazio et al. (1992) demonstrating glutamate's involvement in basal ganglia DA release, light-induced interactions between these two transmitter systems may exist presynaptically. Subsequently, these alterations may trigger biochemical and metabolic changes, such as an up- or down-regulation of transmitter receptor proteins in response to these transmitter levels, which in turn may facilitate the observed biological response. Both DA and glutamate levels were elevated at 24 hours constant light and dark, respectively. Reports from Knapp et al. (1987) demonstrated DA's role in increasing the sensitivity of retinal cells to glutamate excitotoxicity. Thus, elevated DA levels working in concert with elevated glutamate levels may subject the neuronal system to excitoxic damage unless some compensatory mechanism is activated and maintained. Interestingly, KYNA levels, a tryptophan metabolize and endogenous EAA-receptor antagonist, were found to be suppressed at 48 hours of constant light. This suggests a loss of KYNA's neuroprotective antagonistic activity at EAA receptors and a potential increase in EAA excitotoxic susceptibility during light photoperiods. Clearly, an alteration in the natural balance between endogenous DA and EAA receptor ligands may confer varying degrees of susceptibility or protection relative to excitotoxicity in the CNS.

References

1. Difazio M.C. Hollingsworth Z., Young A, B., Penney Jr. J.B. Neurol. 42 (1992) 402-406.
2. Engber Neurosci. 54 (1993) 1051-61.
3. Gamache P., Ryan E., Svendsen C., Murayama K. Acworth I. J. Chromatogr. 614 (1993) 213-220.
4. Knapp A.G. and Dowling J.E. Nature 325 (1987):437-9.
5. Martin-Iverson M. and Yamada N. Brain Res 538 (1991) 310-312.
6. Martin-Iverson M. and Yamada N. Eur. J. Pharm. 215 (1992) 119-125.
7. Stampi C. J. Clin. Pharmacol. 34 (1994) 518-534.

Acknowledgement

This project was supported by NIH grants NS34194, RR03034, and GM08248 and NASA grant NCCW-0083.

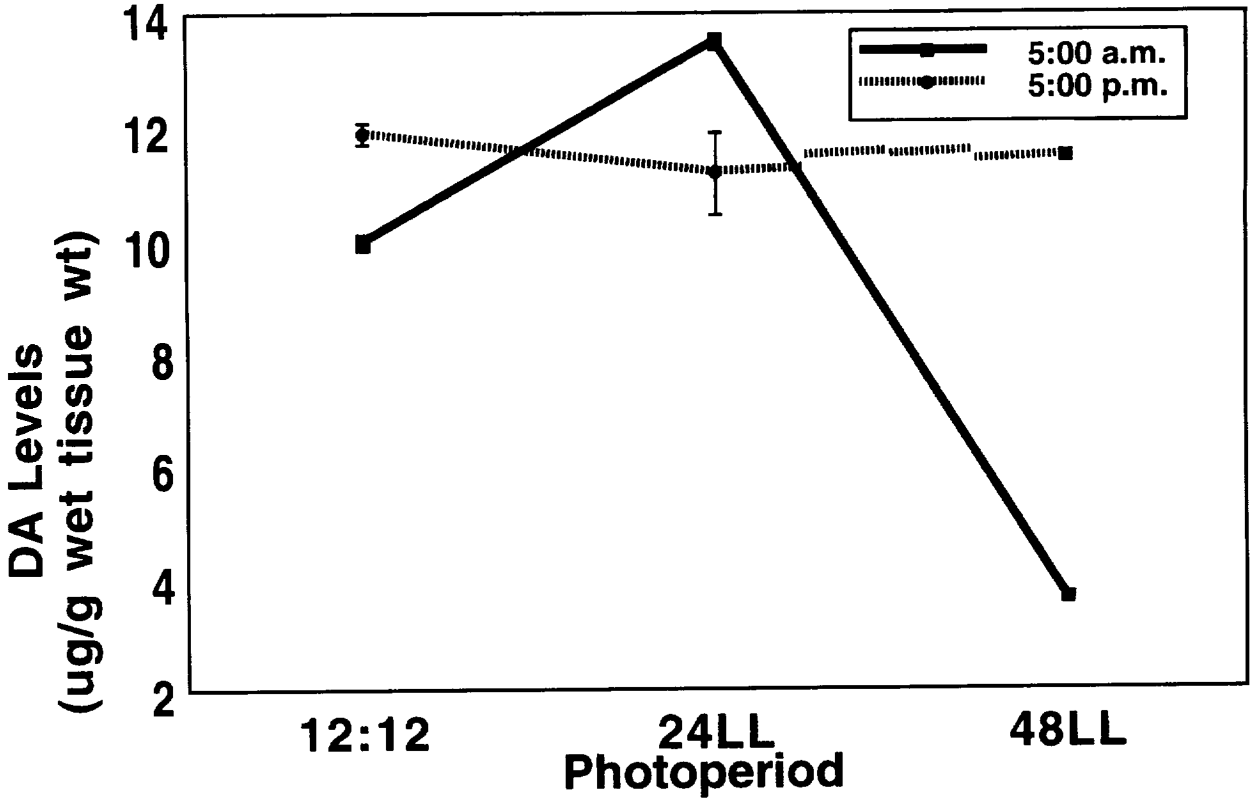


Figure 1 Dopamine levels in rat striatum following exposure to continuous light.

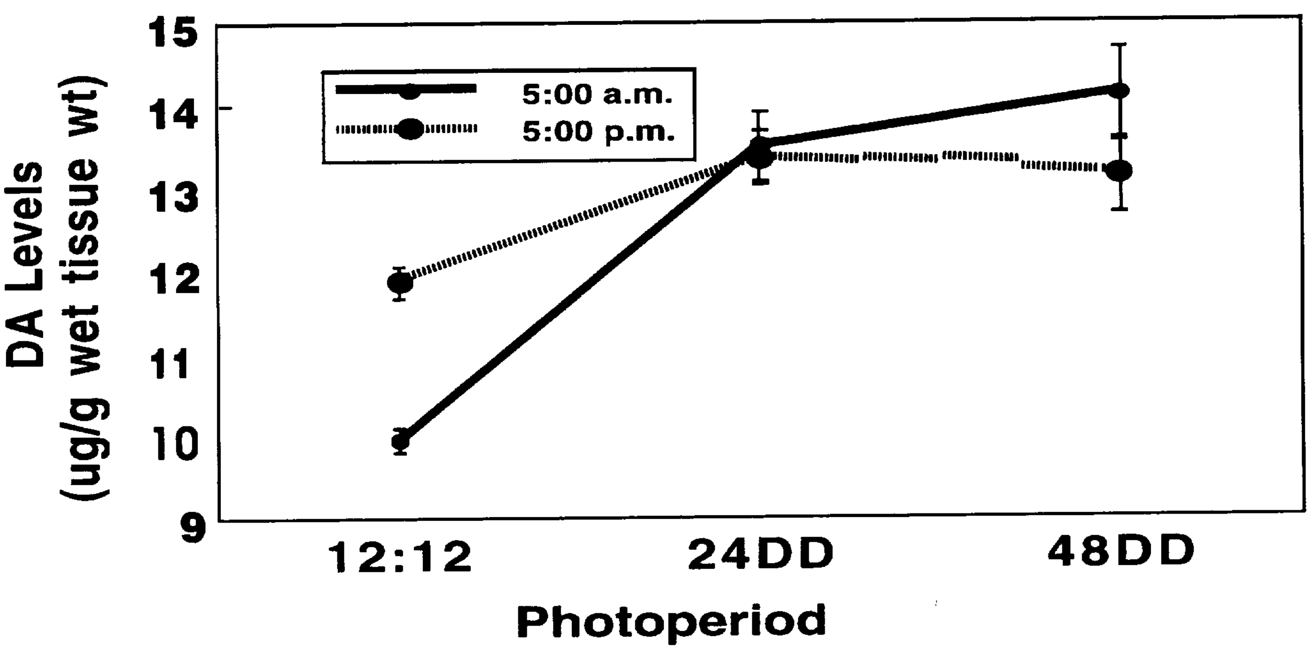


Figure 2. Dopamine **levels in rat** striatum following **exposure to continuous dark.**

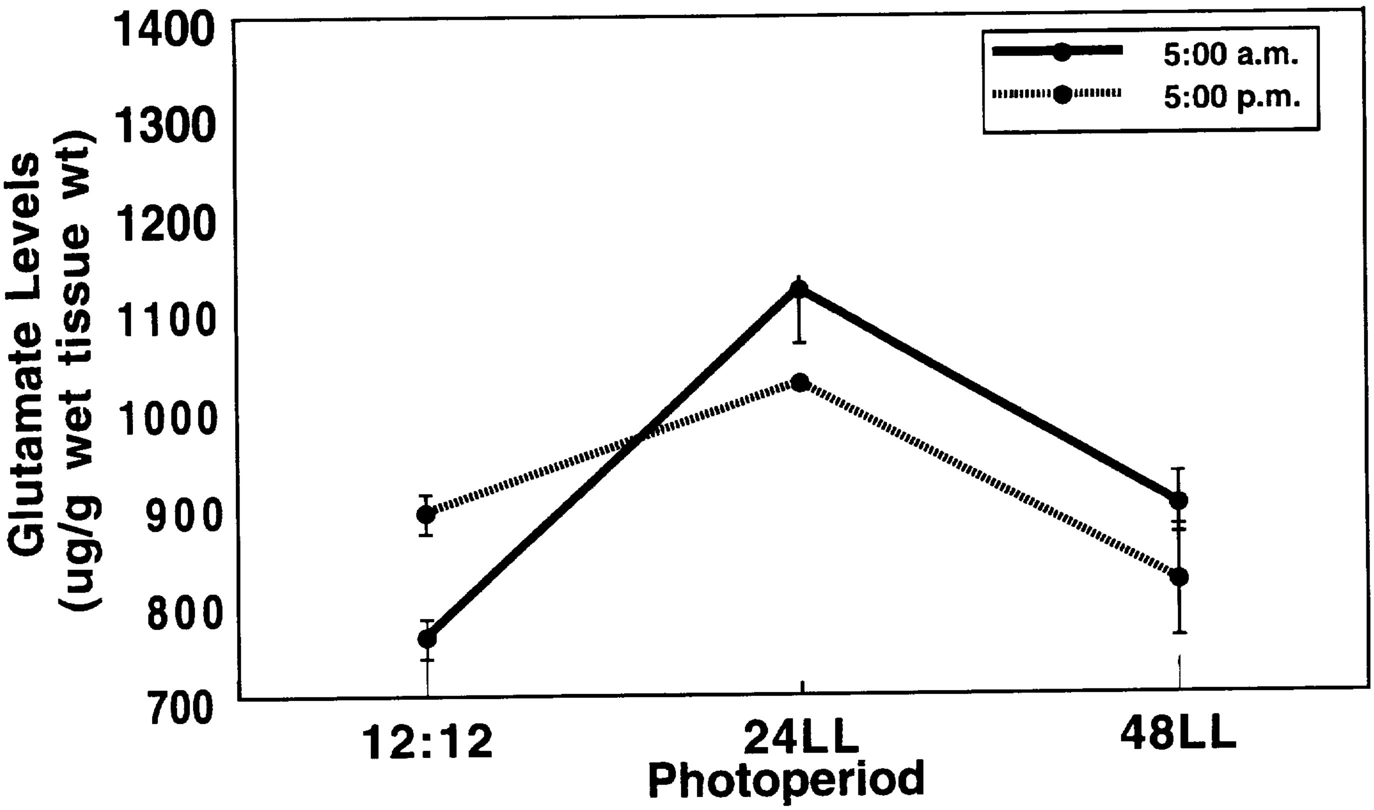


Figure 3. Glutamate levels in rat striatum following exposure to continuous light.

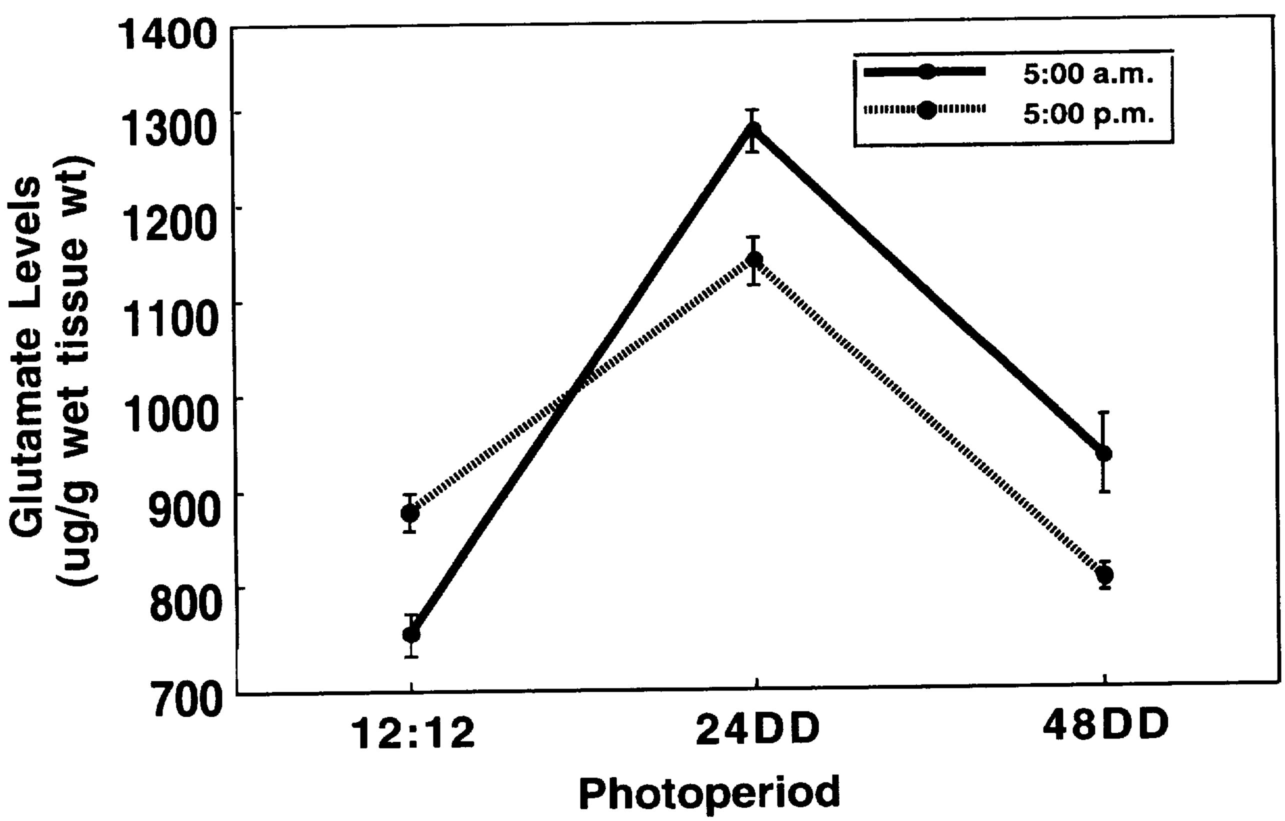


Figure 4. Glutamate Levels in rat striatum following exposure to continuous dark.

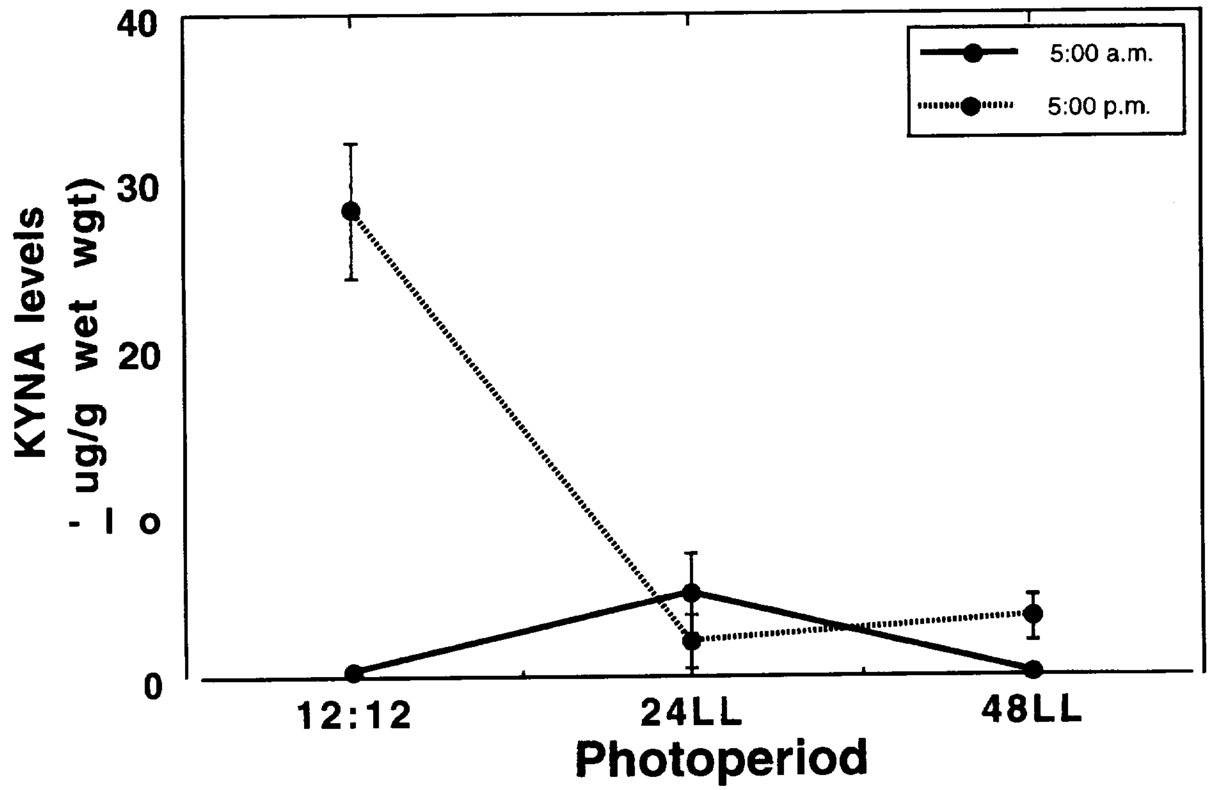


Figure 5. Kynurenic Acid levels in rat striatum following exposure to continuous light.

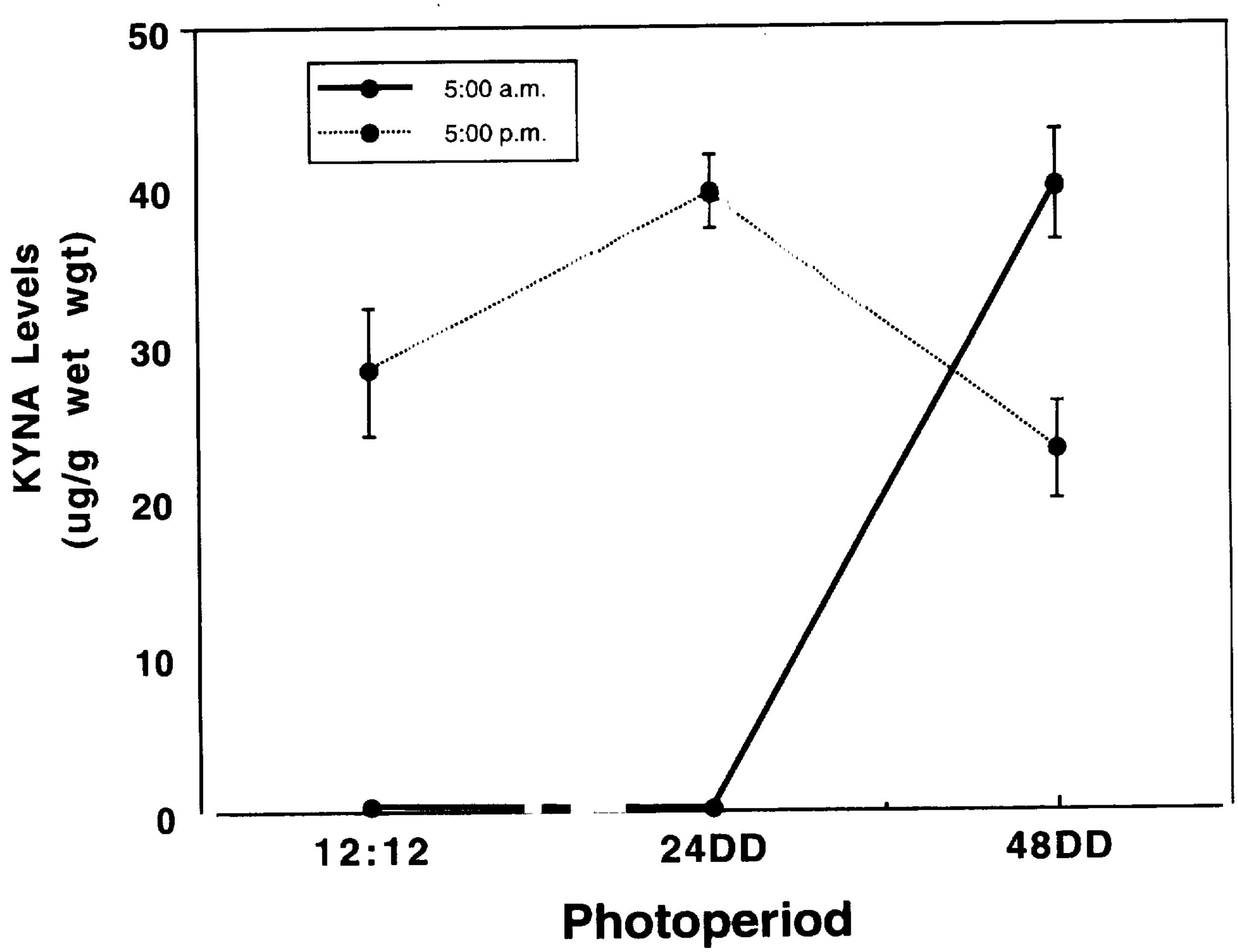


Figure 6. Kynurenic Acid levels in rat striatum following exposure to continuous dark.

NH 5/1

URC97120

An Undergraduate Intern Program at PACES

Scott A. Starks
Pan American Center for Earth & Environmental Studies
University of Texas at El Paso
El Paso, TX 79968
e-mail: sstarks@utep.edu

INTRODUCTION

The University of Texas at El Paso (UTEP) established the Pan American Center for Earth and Environmental Studies (PACES) in 1995 to conduct basic and applied research that contributes to NASA's Mission to Planet Earth [1]. Specifically, PACES provides a repository of remote sensing and other information that supports investigations into an improved understanding of geological, ecological and environmental processes occurring in the southwestern United States and Northern Mexico,

Approximately 85% of UTEP's students come from El Paso County, a fast growing urban region representative of many large cities in the Southwest that have, or will soon have, a majority of their population composed of groups currently underrepresented in the scientific and technical workforce. UTEP's student population has an ethnic distribution (63% Hispanic, 32% Anglo, 3% African American, 1.5% Asian American, and less than 1% Native American) that closely matches the demographics of the region it serves. Thus, UTEP has a mission to serve a multicultural population where minority students comprise the majority.

Most Hispanic students at UTEP are primarily of Mexican origin. A large number are first or second-generation U.S. citizens. Characteristics that unite Hispanic students, in particular those of Mexican-origin, are a strong sense of family loyalty and a belief that all family members are responsible for contributing to the economic stability and well-being of the family. Most of their families are larger in number than the national average, and a variety of generations live together or share considerable resources. Thus, many young people feel an obligation and a desire to go to work at a young age and to continue working while in college, thereby assisting their parents and other family members. Older siblings understand that they have responsibilities to do household chores, to aid their younger siblings economically, and to assist elderly family members. This "work ethic" within the context of family responsibilities is often viewed as being as important as higher education aspirations by both parents and university students. As a consequence, much effort at UTEP has been placed in recent years upon creating opportunities for students to earn income while working on campus as undergraduate research assistants on research projects related to their majors. Work assignments of this sort serve to promote retention efforts, to introduce students with the possibility of graduate study and to develop students professionally. Also, the involvement of students in group-oriented research projects at UTEP which require frequent oral and written reports has been proven beneficial at promoting communication skills of students. Because an estimated 60% of UTEP's students are first-generation college students, many are often uncertain graduate education.

PACES UNDERGRADUATE INTERN PROGRAM

The PACES Center has the goal of preparing well qualified young people to enter careers and graduate programs in areas related to environmental, earth and computing sciences and engineering. In an effort to influence students along these lines, the PACES Center initiated an undergraduate intern program during the summer of 1996. The goals of the program are to: (1) introduce student participants to technologies of interest to NASA in particular those of remote sensing and geographical information systems, (2) to develop research skills among student participants through their involvement in a group-oriented research project, (3) to prepare students

to enter a career in an aerospace related field and (4) to encourage student participants to consider graduate school as an option after completion of their baccalaureate studies.

In light of these goals, ten undergraduate students drawn from engineering and science were selected to form the initial cohort of PACES undergraduate research interns. The students were recruited through an application process based in part upon grades but also on the basis of recommendations from faculty and personal statements written by the students relating to their future plans, including interest in graduate studies. In order to expand the pool of applicants, formal linkages were made with the Alliance for Minority Participation (AMP) for the University of Texas System. Whereas PACES has formal collaborative ties with the University of Texas at San Antonio, one of the selected students was from that institution with the rest from UTEP. All students which were selected were Hispanic. For their participation in the program, the students are rewarded with a monthly stipend, which was paid through funds provided with our grant from NASA.

The underlying technologies of NASA's Mission to Planet Earth of primary interest to the PACES Center are remote sensing and geographical information systems (GIS). The use of remote sensing as a tool for analyses of environmental, cultural, and natural resource management characteristics is well documented [3]. A natural companion to remote sensing as a tool for solving problems associated with land-based processes is the geographical information system (GIS). In addition to representing landscape features, a GIS can be used in management to predict the consequences of a contemplated action, evaluate the results of actions that have been taken, and compare alternative actions. In most basic terms, a GIS is a computerized mapping system for capture, storage, management, analysis and display of spatial and descriptive data.

During the summer, the student interns were provided with a series of lectures on different aspects of remote sensing and geographical information systems. The lectures were held daily in the PACES Center over an eight week period. The textbook by Sabens, [3], and its accompanying laboratory manual were used for instruction. Additionally, the students were provided with instruction in the use of the image processing software package, PCI and a geographical information software system, SPANS Explorer,

The PACES Center provides summer support for eight science and engineering faculty at UTEP who are involved in projects relating to the PACES research mission. Each of these faculty gave an afternoon-long seminar on the research that they were performing and how it related to NASA's MTPE. At these seminars, students were able to recognize how remote sensing could be applied to solve down to earth problems. Additionally, they were able to ask questions and were 'able to observe what a future appointment as a graduate student might be like.

AFRICANIZED HONEY BEE INVESTIGATION

In 1956, the African honey bee was introduced into Brazil. The hope was to create a strain of honey bee which was better suited for the tropical environment of South America and to improve honey production. In 1957, 26 colonies escaped and began interbreeding with the native bees. These new Africanized Honey Bees (AHB's) have a greater tendency to swarm and abscond than domesticated European honey bees. The AHB's have been spreading at a rate of 80 to 500 kilometers annually. At the present time, the AHB has successfully colonized much South and Central America. At the present time, there appears to be no known obstacle that will prevent its colonizing the warmer areas of the United States [4]. In the path of their northward migration, they have reduced honey production and negatively impacted crop pollination, livestock production, tourism and public health. The migration of the AHB's reached the United States in October, 1990 near Hidalgo, Texas in the lower Rio Grande valley. At the present time, AHB sightings have occurred in parts of south and west Texas, southern New Mexico, southern Arizona, the central valley of California.

The impact of AHB's on local economies has been significant. Based on 1981 prices [5], the U. S. beekeeping industry may experience annual losses in the range of $26-$58 million if the AHB

colonizes the Southern and Southwestern U.S. and causes the same types of problems it has caused in South and Central America. In some areas, beekeeping and honey production declined because as many as 80% of the amateur beekeepers abandoned their colonies, up to 20% of the commercial beekeepers quit beekeeping, and hone y production by the remaining colonies declined. Although the AHB is implicated in changes in honey production, other factors including weather, destruction of forests and increased pesticide use (especially on cotton) may have contributed to reduced honey production.

At the Spring Advisory Board meeting for the PACES Center, research personnel from the Center for Health Applications of Aerospace and Related Technologies at NASA Ames Research Center suggested that it would be interesting to investigate the northward migration of Africanized Honey Bees (AHB's) using remote sensing and geographical information systems as tools. From this recommendation came the centerpiece group research project for the PACES undergraduate interns. As a consequence, the summer interns were organized into a research team with the goal to develop a GIS to track the northward migration of the AHB and to better understand the interaction of climatic, ecological and environmental factors influencing their movement.

The PACES Center acquired a software package for GIS development, SPANS Explorer, which is distributed by TYDAC Technologies, Inc. The SPANS Explorer software package has several formats for data layers and methods for analysis well suited for the AHB migration study. As background, at the core of any GIS is a base map. Information can then be added to complement the base map by overlaying data layers.

The attribute data that is being used in the development of the AHB migration GIS can be grouped into the categories of climatic, land usage and bee sighting data. The data in each of these categories are incorporated into the GIS as data layers. The first category, climatic data, consists of the number of frost free days, the average monthly temperature and amount of precipitation. This data is compiled on a county-wide basis. All three of these factors have been reported as important factors influencing the migration of AHB's [5]. The climatic data was acquired form the National Weather Service and the National Climate Data Center. The second category of data, land usage, is being assembled on the basis of information being acquired from State Departments of Agriculture and Commerce. The third category of data, bee sightings, provides the most important information of likely bee migration.

Information regarding the characteristics and spatial distribution of the Earth's land cover is critical to understanding AHB migration. During the past decade, substantial progress has been made in using National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) data for land cover characterization [6]. AVHRR data have only moderate spatial resolution (1 km) when compared, for example, to Landsat's 30 m resolution for thematic mapper or SPOT's 20 m for multispectral and 10 m for panchromatic data. However, the AVHRR data is collected more frequently with virtually entire global coverage twice each day. The high frequency of coverage enhances the likelihood that cloud-free observations can be obtained for specific temporal window, and makes it possible to monitor change in land cover conditions, over short periods, such as a growing season.

At the present time, remote sensing data obtained from the AVHRR sensor is also being incorporated into the GIS through a data layer which stores vegetation greenness. Greenness is most often quantified using a vegetation index, commonly the Normalized Difference Vegetation Index (NDVI) [7] which can be calculated using the AVHRR sensor suite

In our investigation of AHB migration, AVHRR data is being used to characterize land cover in our study area which includes much of the southern border region of the United States with Mexico. To date, we have imported hi-weekly NDVI calculations back to the year 1989. We are using time series of NDVI readings to stratify regions of vegetated land cover from barren land. We next intend to perform an unsupervised learning procedure to extract clusters of the various land cover types in our study region.

SUMMARY AND FUTURE WORK

The progress made by the PACES Undergraduate Interns during the summer period was impressive, The group was able to develop a rudimentary GIS with attribute data relating to climate and bee sightings. At the end of the summer, PACES sponsored a trip to NASA Ames Research Center to enable the students to present their research results. Two group presentations were made: one stressing the biological characterization of the AHB's and the other emphasizing the software aspects of developing the GIS. While at Ames, the interns were able to interact with numerous scientists and engineers as well as to visit numerous research facilities. In addition, the students were able to gather information about graduate education opportunities in the San Francisco area.

Following the summer, all interns continue to work on various aspects of the project as undergraduate research assistants. At the present time, correlation studies which are based on observations of the standard deviation of NDVI measurements from the AVHRR data are being conducted. As an adjunct to the project, the interns have accepted the challenge of community service by developing plans for the construction of a bee observatory. The bee observatory will house actual bee colonies and will include information on the habits and behavior of bees. Plans call for the building of two bee observatories. One will be located on the grounds of the Chamizal National Monument located on the international border in El Paso and the other will be sited on the UTEP campus. It is hoped that the bee observatories will serve to educate the public on bees and their value to our society and the economy. They will also be utilized in a variety of science outreach programs directed at young students.

At the present time, we envision a series of student publications to result from this program [8]-[10]. Also, one student will soon complete a senior Honors Thesis on work completed through the program. The experience gained from the preparation and presentation of research papers at professional meetings and to groups of professionals has had a very beneficial impact on all student participants.

ACKNOWLEDGEMENTS

We wish to thank NASA for its support of this project through the grant NCCW-0089. Also, we wish to extend gratitude to Byron Wood and Louisa Beck of the Center for Health Applications of Aerospace and Related Technologies at NASA Ames Research Center for their assistance in the formulation of this project and Pablo Arenaz and James Smash of the University of Texas System Alliance for Minority Participation for helping identify and recruit undergraduate student interns.

REFERENCES

[1] Starks, S.A, D.E. Cooke and G.R. Keller, "Earth and Environmental Remote Sensing at PACES, " to appear in Geocarto, Int'l.

[2] Quattrochi, D.A. and R.E. Pelletier, "Remote Sensing for Analysis of Landscapes: An Introduction, " in Quantitative Methods in Landscape Ecology, M.G. Turner and R.H. Gardner, eds., New York: Springer Verlag, 1991.

[3] Sabens, F.F., Remote Sensing: Principles and Interpretation, New York: Freeman, 1987.

[4] Gary, N. E., H.V. Daly, S. Locke and M. Race, "The Africanized Honey Bee, Ahead of Schedule," California Agriculture, pp. 4-7, November 1985.

[5] McDowell, R., "The Africanized Honey Bee in the United States," U.S. Department of Agriculture, Agricultural Economic Research Report No. 519, 1984.

[6] Loveland, T. R., J.W. Merchant, D.O. Ohlen, and J.F. Brown, "Development of a Land-Cover Characteristics Database for the Conterminous U.S.," Photogrammetric Engineering & Remote Sensing, Vol. 57, No. 11, pp. 1453-1463, November 1991.

[7] Lillesand, T.M. and R.W. Kiefer, Remote Sensing and Image Interpretation, New York: John Wiley, 1987.

[8] Navarro, Hector, "An Investigation of the Migration of Africanized Honey Bees into the Southern United States," Proc. of the 1997 NASA URC-TC '97 National Technical Conference on Education, Aeronautics, Space, Autonomy, Earth, and Environment, Albuquerque, NM, Feb. 1997.

[9] Ward, C. *et.al.*, "The Integration of Geographical Information Systems and Remotely Sensed Data to Track and Predict the Migration Path of the Africanized Honey Bee," Proc. of the 1997 NASA URC-TC ' 97 National Conference on Education, Aeronautics, Space, Autonomy, Earth and Environment, Albuquerque, NM, Feb. 1997.

[10] Bravo-Flores, J. and H. Navarro, "An Investigation into the Migration of Africanized Honey Bees," to be presented at the Winter Meeting of the American Association of Physics Teachers, Phoenix, AZ, 1997.

Page intentionally left blank

URC97121

5121/27

Crystalline Colloidal Arrays in Polymer Matrices

Hari B. **Sunkara***, B. G. Penn, D. O. Frazier, N. **Ramachandran**‡
NASA Marshall Space Flight Center, Space Sciences Laboratory, **Huntsville, AL** 35812
•National Research **Council,** ‡ Universities Space Research Association

Introduction

Crystalline Colloidal Arrays (**CCA, also known** as **colloidal crystals),** compose of aqueous or **nonaqueous** dispersions of self-assembled **nanosized polymer colloidal** spheres, are emerging toward the development of advanced optical devices for technological applications. **The** spontaneous **self assembly** of polymer spheres in a **dielectric** medium results **from** the **electrostatic** repulsive **interaction** between particles of uniform size and charge distribution. In a way similar to atomic **crystals** that **diffract** X-rays, CCA dispersions in thin quartz cells selectively and efficiently Bragg diffract the incident **visible light.** The reason for this diffraction is because the **lattice** (body or face centered cubic) spacing is on the order of the wavelength of **visible light.** Unlike the atomic **crystals** that **diffract** a **fixed** wavelength, **colloidal crystals** in principle, depending on the panicle size, particle number and charge density, can **diffract** *W*, **Vis** or **IR light.** Therefore, the **CCA** dispersions can be used as laser filters. Besides, the **diffraction intensity** depends on the refractive index mismatch between **polymer** spheres and dielectric medium; therefore, it is possible to modulate incident light intensities by manipulating the index of either the spheres or the **medium.**

Our interest in CCA is in the fabrication of all-optical devices such as optical switches, limiters, and spatial **light** modulators for optical signal processing. The two major requirements from a materials standpoint are the incorporation of **suitable** nonlinear optical materials (NLO) into polymer spheres which will allow us to alter the **refractive** index of the spheres by intense laser radiation, and preparation of solid **CCA** filters which can resist laser damage. The fabrication of **solid** composite filters not only has the advantage that the **films** are easier to handle, but also the arrays in **solid** films are more robust than in liquid media. In this paper, we report the **photopolymerization** process used to trap CCA in polymer matrices, the factors which affect the optical **diffraction** qualities of resulting polymer **films,** and methods **to** improve the efficiencies of solid **optical** filters. Before this, we also present the experimental demonstration, of controlling the optical **diffraction** intensities **from** aqueous **CCA** dispersions by varying the temperature, which establishes the **feasibility** of fabricating all-optical switching devices with nonlinear **periodic** array Structures.

Temperature **Switchable Optical** Filters

Figure l shows the dramatic optical **diffraction** responses of aqueous CCA of **poly(N-isopropy lacrylamide) microgel** (PNIPAM) spheres as a **function** of temperature. [1,2] The CCA barely **diffracted** the incident light at IO 'C, but **diffracted** the **light** very effectively at 40 'C, and thus, this CCA dispersion acts as an optical switch. The **optical** responses of CCA with temperature are **highly** reversible, and no hysteresis was noticed even after several rapid heating and cooling cycles of the samples.

The **PNIPAM** gel is an interesting **thermosensitive** polymer which exhibits a reversible volume change in aqueous medium at 32 'C, causing many of its **physico-chemical** properties to change concurrently. Below the phase transition temperature, the low cross linked **microgel** spheres are highly swollen and have a refractive index value that is close to water. Above this temperature, the spheres are collapsed and have a refractive index close to the polymer. The decrease in **particle** size and increase in the refractive index of the spheres with increase in temperature have opposite effects on the particle scattering intensity. The observed increase in diffraction intensity indicates that the contribution from the changes in the refractive index dominates more than the particle size change in altering the optical properties.

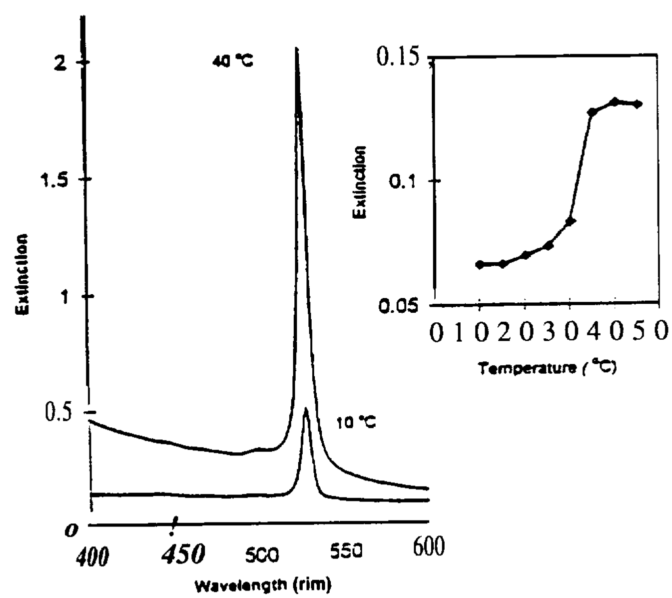


Figure l. **Optical** diffraction intensity (left) from aqueous ordered dispersion **of 3%** crosslinked poly(N-isopropy lacrylamide) hydrogel spheres as a *function* of temperature. Light scattering from disordered dispersion of very low concentration of spheres as a function of temperature (right).

The scattering intensity from a very dilute colloidal dispersion (0.01 wt% solids) as a function of temperature is shown in the inset of Figure 1. Between the two extreme temperatures studied, the scattering intensity increased only by a factor of 2, however, the diffract ion intensity increased by a factor of 6. These experiments indicate that it is possible to create a laser induced optical switch or limiter with CCA incorporated with NLO materials.

Colloidal crystals suspended in liquid media are very sensitive to ionic impurities and, as well, to any external weak forces such as shear[3], heat, gravity[4,5] and electric fields that can melt these crystals. The fragility of the crystallites is due to their low bulk and shear moduli. Therefore, it is essential to convert these delicate structures into robust systems in order *to* use them in optical device applications. The recently developed photopolymerization methodology of traping the macroscopic ordered structures in solid polymer matrices is very successful in overcoming the aforementioned problems.[6-13] However, at the present time, the development of these polymer nanocomposite filters is not fully matured yet, and suffers from a lack of reproducibility of optical diffraction properties either from one polymerized film to another or from the prepolymerized fluid film to its corresponding polymerized film. The following section analyzes the method of preparation of solid filters and their optical properties to understand the dynamics of crystallite during the photoinitiated polymerization processes.

Crystalline Colloidal Arrays in *Acrylic* Polymer Matrices

To manufacture solid nanocomposite laser filters from liquid CCA dispersions, we initially modified the surface of colloidal silica spheres by treating them with the silane coupling agent, 3-(trimethoxy)silylpropyl methacrylate (TPM). Transferred the TPM coated silica spheres into either methyl acrylate (MA) or methyl methacrylate (MMA) monomer, and allowed the spheres to self assemble before irradiating the monomeric dispersions.[6-10] The photopolymerization experiments were conducted at ambient temperature for 4 h using a medium pressure mercury arc lamp. Figure 2 shows the optical diffraction from liquid and solid crystalline colloidal dispersions. A very large shift of the Bragg diffraction peak and a much wider bandwidth is obvious for the polymerized film. The diffracted wavelength shifted from 554 to 490 nm which corresponds to a 15°/0 decrease in *d* spacing or a 45°/0 compression of the original lattice. The peak bandwidth increased from 4 nm to about 15-20 nm in the polymerized film. The lattice compression in the polymerized film has been attributed mostly to the volume shrinkage of the host matrix. Based on the observation that only the thickness (but not length and width) of the polymerized film decreased, it has been proposed that the crystalline structure transforms from a face centered cube to a rhombohedral during the photopolymeri-zation process.[9-11]

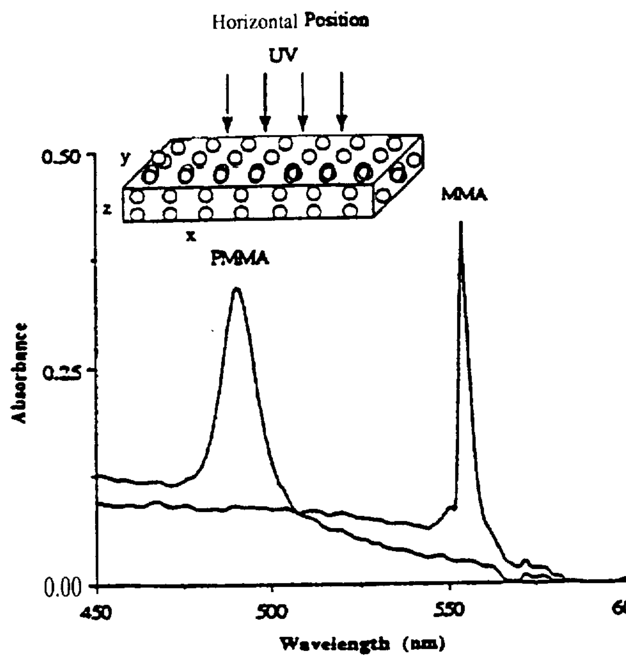


Figure 2. Bragg diffraction from arrays of silica spheres (35 wl'??) in methyl methacrylate before and after photopolymerization. The photocell (length = 70mm, width = 20mm, and thickness " 0.264 mm) was placed horizontally on a lab jack.

This model raises the following questions: why does the volume of the monomeric dispersion decrease only in thickness during polymerization? For example, the volume fraction of MMA used in the dispersion is 0.805 and hence a 17% decrease in volume (21% for neat MMA) can be expected after the solidification process. The decrease in volume is due to the loss of the bulky π electron cloud which is perpendicular to the carbon-carbon double bond. Since the volume change of the monomer matrix occurs at the molecular level, the volume should shrink uniformly in all three dimensions with a 5.6% decrease in each dimension. However, the host volume in the polymerized film decreased profoundly in one dimension, that is in thickness, by about 14%.[9] Secondly it is not clear how the microscopic volume changes during polymerization affects the lattice dimensions and the lattice structure during polymerization. On the other hand, the effect of macroscopic volume shrinkage of the host polymer network on crystal compression, when a temperature sensitive poly(N-isopropyl acrylamide) hydrogel network embedded with arrays of polystyrene latex spheres is exposed to heat, has been examined. The volume of the gel network shrunk isotropically with increase in temperature, and the network compressed the crystal lattice uniformly. 'Thirdly, this model does not take into account the dynamics of colloidal crystals. It is well known that the spheres in a lattice vibrate continuously from their equilibrium positions, and colloidal crystals dynamically respond to external radiation stimulus.[14-16]

Dynamics of Colloidal Crystals

During the photopolymerization of silica-MMA dispersions, the dynamics of colloidal crystals changes dramatically, not because of the stress of volume shrinkage, but due to the changes in the dielectric properties of the host matrix. As the monomer converts to polymer, the dielectric property of the medium changes for two reasons. Firstly, the polarizability of vinyl monomer is greater than the monomer unit in the corresponding polymer and secondly the heat liberated during reaction decreases the dielectric constant. The dielectric constant of PMMA (2.60 at 1 MHz) is less than MMA (6.32 at 1 MHz) by a factor of 2.4. This decrease in dielectric constant of the host decreases the electrostatic repulsive interaction between the spheres as a result of counter-ion association. This causes the colloidal crystals to shrink during polymerization. The interparticle interaction potential (U) between two spheres of radius a at a distance r is a strong function of the dielectric constant (ε) of the medium, and is shown in Eq (1)[15]

$$U(r) = \frac{(Ze)^2}{\varepsilon}\left(\frac{\exp(ka)}{1 + ka}\right)^2 \frac{\exp(-kr)}{r} \tag{1}$$

where Ze is the particle charge, and the inverse Debye screening length k is given by

$$k^2 = \quad \left(n_p Z + n_i\right) \tag{2}$$

where n_p is the particle concentration, n_i is the ionic impurity concentration, k_B is the Boltzmann constant and T is the temperature.

The effect of polarizability of dielectric medium on colloidal crystals of TPM silica has been demonstrated by Philipse and Vrij.[17] The number of charges on a TPM silica sphere decreased from -500 to 90 as the dielectric constant of the medium was changed from a pure ethanol (ε=25) to a mixture of toluene/ethanol, 70/30 v/v (ε=10). This led to a decrease in inverse Debye length from 100 nm to about 50 nm. The effect of local heating of colloidal crystals using a laser beam on lattice compression has been investigated by Asher *et al.*[15,16] The localized compression of the dyed colloidal crystals has been attributed to reduced electrostatic interaction between the spheres which results from the temperature dependent dielectric constant of the medium. These experiments clearly support our assumption that the change in dielectric constant of the matrix is responsible for the lattice shrinkage during polymerization. The accompanied shrinkage in volume of the host minimizes the disorganization of sphere arrays.

The broader bandwidths resulting from a decrease in the degree of orderness of spheres, heterogeneities of optical diffraction properties within a polymerized film and the decrease in film thickness are ascribed to the effect of gravity induced convection and sedimentation. Before arguing the effect of convection on organized colloidal spheres in monomer matrix during polymerization processes, we discuss the origins of convection in silica-MMA dispersion, and the factors which influence the convective flows most.

Convective Instabilities

It is well known that gravity induced convective flows arise in an unstirred system whenever density gradients exist.[18] There may be several variations that can create density gradients: the presence of more than one phase, material phase transitions from one state to another, and differences in either temperature or concentration.

All of the variations mentioned above which can produce density differences do exist when a colloidal dispersion is subjected to irradiation. As UV radiation penetrates into the methyl methacrylate dispersion in which the arrays of submicron silica spheres are suspended, the light attenuates gradually in the direction of propagation because of the absorption of light by the photoinitiator and monomer molecules present in the dispersion. Further, the light attenuates due to scattering by the colloidal spheres which are in the of 100-150 nm diameter range and occupy about 20% of the total volume. In contrast to the absorption of light by the photoinitiator and MMA molecules, the photon scattering by the colloidal particles produce no heat in the dispersion. The intensity losses due to absorption and scattering result in an intensity gradient along the direction of light propagation in the dispersion. Because of this intensity gradient in the dispersion, the polymerization rate, which is intensity dependent, varies in the dispersion. Hence, the reaction rate at the dispersion layers where the light enters would be higher and decreases gradually in the medium. Since the addition polymerization reaction is exothermic (the heat of polymerization of MMA is 13.0 kcal/mole), the liberated heat causes an added thermal gradient in the dispersion due to lack of thermodynamic equilibrium. Additionally, the glass/quartz cell containers used to fill the colloidal array dispersions also absorb the UV radiation and cause thermal gradients in the dispersion. These thermal gradients in

the dispersion generate density gradients ($\Delta\rho_{thermal}$ is negative), and the density gradients under the influence of gravity induces convective fluid motions.

Furthermore, a large decrease in the partial molar volume of the matrix and concentration changes of the monomer during polymerization induce density gradients ($\Delta\rho_{monomer}$ is positive) in the dispersion.[19] In addition to density gradients created by the matrix monomer during the reaction, density gradients exist even before the reaction starts due to the presence of denser colloidal silica spheres (1 .79 g/cm^3) dispersed in the lighter MMA (0.94 g/cm^3). As mentioned, the sphere concentration also changes due to reduced interparticle interactions, and heterogeneities in the reaction rate across the dispersion due to intensity gradients cause density gradients ($\Delta\rho_{spheres}$ is positive). Therefore, a net density change ($\Delta\rho$) during the polymerization of monomeric dispersion is

$$\Delta\rho = \Delta\rho_{thermal} + \Delta\rho_{monomer} + \Delta\rho_{spheres} \qquad (3)$$

Since the photopolymerization experiments are conducted at ambient temperatures, we presume that the density gradients which arise from the thermal effect is smaller than the density gradients resulting from concentration changes of monomer and spheres (solutal convection), and the volume changes of the monomer matrix (double diffusive convection). The reacted top layers in the horizontally held dispersion become denser because of the positive $\Delta\rho$ and are positioned above the less dense unreacted fluid layers. Therefore, under the influence of gravity, these dense layers sink to the bottom and a decrease in film thickness results after polymerization. This decrease in fluid film thickness during polymerization can create free surfaces and, consequently Marangoni convection may develop in addition to buoyancy effects due to the presence of density gradients in these systems.[20] The magnitude of fluid motions may be small due to small cell container dimensions, but large enough to deform the de licate crystalline structures. The following experimental observations also indicate that the gravity driven heat and mass transport play a crucial roles in the photopolymerization of monomeric dispersions in thin ceils.

The silica-MMA dispersion polymerized in the horizontal configuration (Fig.2) showed optical diffraction whereas the dispersion polymerized in the vertical configuration did not. Since no other experimental parameters are changed except the orientation of the cell, which can affect neither the heat of polymerization nor the volume shrinkage of the host, the observed phenomenon has to be gravity related. The disappearance of Bragg diffraction from the solidified films is a result of either complete destruction of ordered arrays of spheres or disorientation of the lattice planes of the crystallite. Crandall and Williams,[4] and Kesavamoorthy and Arora[5] have reported that the gravitational force significantly affects the interacting colloidal polystyrene spheres (100 nm in diameter and density of 1.05g/cm^3)) dispersed in aqueous medium and compresses the lattice. Therefore, we initially thought that the disappearance of the Bragg diffraction peak from the silica-PMMA dispersion, which was kept in the vertical position just before polymerization, is due to the gravitational sedimentation of the colloidal crystals. The crystallites of silica spheres, which were grown for a few days with the cell container in the horizontal position, are in gravitational equilibrium because the rate of sedimentation equilibrium is faster than the rate of crystallization. Upon tilting the position of the ceil container to vertical orientation, these crystallite, under the influence of gravity, tend to reach a new sedimentation equilibrium, and can deform. However, the Bragg diffraction peak was not destroyed in the unpolymerized dispersion, although the peak position and shape change, when the fluid sample was held verticality in a spectrophotometer for 3-4 h, the duration of the photopolymerization process. These experimental results suggest that the absence of diffraction from the vertically held polymerized sample is not simply due to sedimentation of crystallites under the influence of gravity.

It appears that the polymerization process makes the fluid more unstable in a vertically oriented photocell than in one horizontally orientated. In the vertical configuration, as mentioned above, the gravitational force acts on colloidal crystals. In addition to this, density gradients in the dispersion, created by the intensity dependent reaction, induce convect ive flows. Considering the fact that the reaction liberates heat, the density of dispersion layers which are close to the radiation source become lower and gravity, therefore, induces movement of these layers. On the other hand, these layers become denser if the effect of solutal and double diffusive convection dominate thermal convection. Again, in this case, gravity drives the motion of the fluid layers. In any case, the fluid in the vertical configuration is unstable and the gravitational force can randomize the sphere arrays during polymerization. However in the case of the horizontal configuration, the fluid layers which are close to the radiation source are on top, and these become less dense due to positive AT. Since these layers stay over the dense bottom layers, gravity may not induce convection, and thus the dispersion in this configuration is stably stratified. The top layers become denser due to positive $\Delta\rho$ as mentioned above, and sink to the bottom by the acceleration due to gravity. The sedimentation of dense layers may not randomize the sphere packing, but lattice dimensions certainly change. That is why the polymerized film obtained in the horizontal position diffracted the incident light (Figure 2).

In another experiment, in spite of keeping the photochemical cell in the vertical position, the Bragg diffraction has been observed from the solidified poly(MMA-*co*-HEMA) film (Figure 3).[8] As stated earlier, the vertical configuration is highly unstable because of buoyancy induced fluid fluctuations. Such motions in general are opposed by viscous drag forces. The HEMA (2-hydroxyethyl methacrylate) is more viscous and denser than MMA. Therefore the bulk viscosity of the monomer mixture (MMA/HEMA; 65/35 wt%) is higher than neat MMA and hence the high viscous matrix stabilizes the dispersion against convection and thereby minimizes the mass transport as observed with silica-PMMA film. As a result, the periodic arrays of silica spheres are not disturbed completely after polymerization. However, the crystalline lattice is still compressed during the reaction and this is because of reduced interparticle interactions as stated earlier. The attempts to self-assemble the colloidal silica spheres in neat HEMA are not successful because of the high viscous matrix which reduces the rate of crystal growth.

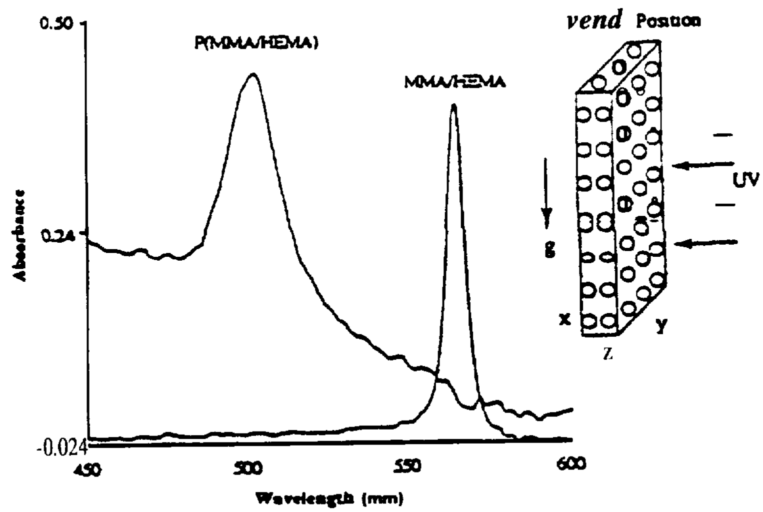


Figure 3. Bragg diffraction from silica-(MMA-co-HEMA) dispersion before and after polymerization. The photocell was placed vertically just before polymerization.

One of the major components in the dispersion which could lead to large convective instabilities during the polymerization process is the photoinitiator. if both initiator concentration&d its extinction coefficient are high in the dispersion, intensity gradients can produce large scale buoyant forces. The self-screening effect by the photoinitiator molecules can not be eliminated but maybe reduced by carefully choosing an optimum concentration of photoinitiator for a given thickness of the dispersion.[21] The initial concentration of the photoinitiator, 2,2-dimethoxy-2-phenyl acetophenone, employed in a 0.26 mm thick dispersion is I wt%. However, later investigations study the effect of varying this photoinitiator concentration (from 0.2 to 5.0 wt%) on optical diffraction properties of polymethyl acrylate composite films.[9] The diffraction peak bandwidths are narrower at 0.2 wt% which suggest that the periodic arrays are less disturbed at these levels. These results are not surprising because the lower the photoinitiator concentration, the smaller the intensity gradients in the dispersion, which thus minimizes the density gradients. The low initiator concentrations favor minimal convection at the expense of slow reaction rates.

Conclusion

We have identified several factors which influence the dynamics of colloidal crystals during the photoinitiated bulk polymerization process. To design better nonocomposite laser filters, the lattice compression, volume shrinkage of the monomer matrix, and the gravity induced convection and sedimentation have to be minimized. Further, understanding the effect of gravity on colloidal crystals during polymerization is essential. Microgravity provides a unique convection-free and sedimentation-tie environment. Processing of materials in such an environment could lead to a better understanding of the effect of gravity on these processes.

References

(1) Weissman, J. M.; Sunkara, H. B.; Tse, A. S.; Asher, S. A. *Science 1996, 274 (Nov. 8).*
(2) Sunkara, H. B.; Weissman, J. M.; Penn, B. G.; Frazier, D. O.; Asher, S. A. *ACS Polym. Preprint 1996, 37, 453.*
(3) Imhof, A.; van Blaaderen, A.; Dhont, J. K. G. *Langmuir 1994, IO, 3477.*
(4) Crandall, R. S.; Williams, R. *Science 1977, 198, 293.*
(5) Kesavamoorthy, R. ; Arora A. K. *J. Phys. A 1985, 18, 3389.*
(6) Sunkara, H. B.; Jethmalani, J. M.; Ford, W. T. *Chem. Mater. 1994, 6, 362.*
(7) Sunkara, H. B.; Jethmalani, J. M.; Ford, W. T. *ACS Polym. Mater. Sci. Eng. Preprint 1994, 70, 274.*
(8) Sunkara, H. B.; Jethmalani, J. M.; Ford, W. T. in *Hybrid Organic-Inorganic Composites,* Eds. Mark, J. E.; Lee, C. Y-C.; Bianconi, P. A. *ACS Symp. Ser.* 1995, 585, 181.
(9) Jethmalani, J. M.; Ford, W. T. *Chem. Mater.* 1996, *8, 2138.*
(10) Jethmalani, J. M.; Sunkara, H. B.; Ford, W. T.; Willoughby, S. L.; Ackerson, B. A. (submitted [o Langmuir).

(11) **Jethmalani,** J. M.; Ford, W. T.; **Beaucage,** G. (submitted to **Langmuir).**
(12) Panzer, H. P. ; **Magliocco,** L. G.; Cohen, M. L.; Yen, W. S. US Patent 5338492, 1994.
(13) **Asher,** S. A.; **Holtz,** J.; **Liu,** L.; **Wu,** Z. ***J. Am. Chem. Sot. 1994, 116, 4997.***
(14) **Malcuit,** M. S.; Herbert, C. J. ***Acts Physics Polonica .4,1994,86, 127.***
(15) **Rundquist,** P. A.; **Kesavamoorthy,** R.; **Jagannathan,** S.; **Asher,** S. A. ***J. Chem. Phys.*** 1991,95,8546.
(16) **Rundquist,** P. A.; **Jagannathan,** S. ; **Kesavamoorthy,** R.; **Brnardic,** C.; **Xu,** S.; **Asher,** S. A. ***J. Chem. Phys.*** 1991, 94, 711.
(17) **Philipse,** A. P.; **Vrij,** A. ***J. Chem. Phys. 1988, 88,6459.***
(18) Turner, J. S. Buoyancy ***Effects in Fluids,*** Cambridge University Press, Cambridge, 1973.
(19) **Pojman,** J. A.; Epstein, I. R. ***J. Phys. Chem.*** 1990,94,4966.
(20) Winters, K. H.; '1%. **Plesser, Cliffe,** K. A. ***Physics 1988,* 29D, *387.***
(21) Bush, R. W.; **Ketley,** A. D.; Morgan, C. R.; **Whitt,** D. G. ***J. Radiat. Curing 1980,7,20.***

URC97122 /34

Non-Intrusive Optical Diagnostic Methods for Flowfield Characterization

NASA Grant No. NAGW- 1-2929

Bagher M Tabibi, Charles A. Terrell, Darrell Spraggins, Ja H. Lee

Department of Physics
Research Center for Optical Physics
Hampton University
Hampton, VA 23668

and

Leonard M. Weinstein
NASA LaRC
Hampton, VA 23681

SUMMARY/OVERVIEW

Non-intrusive optical diagnostic techniques such as Electron Beam Fluorescence (EBF), Laser-Induced Fluorescence (LIF), and Focusing Schlieren (FS) have been setup for high-speed flow characterization and large flowfield visualization, respectively. Fluorescence emission from the First Negative band of N_2^+ with the (0,0) vibration transition (at λ =391.44 nm) was obtained using the EBF technique and a quenching rate of N_2^{+*} molecules by argon gas was reported'. A very high sensitivity FS system was built and applied in the High-Speed Flow Generator (HFG) at NASA LaRC. A LIF system is available at the Advanced Propulsion Laboratory (APL) on campus and a plume exhaust velocity measurement, measuring the Doppler shift from λ = 728.7 nm of argon gas, is under way.

INTRODUCTION

Advancement of non-intrusive optical diagnostic techniques has been a major area of research since the founding of the Research Center for Optical Physics (RCOP) at Hampton University. Many non-intrusive optical techniques have been developed and successfully applied worldwide in high-speed flow facilities and other areas of research during the past two decades. Among them, Electron Beam Fluorescence (EBF), Laser-Induced Fluorescence (LIF), Rayleigh/ Raman Scattering (RRS), Coherent Anti-Raman Scattering (CARS), Laser Doppler Velocimetry[6] (LDV), and Focusing Schlieren[7] (FS) were commonly chosen by researchers because of their wide range of applications.

Our goals for the development of several non-intrusive techniques (namely EBF, LIF, and FS) at RCOP were two-fold: first, to apply these techniques in the aerodynamic wind tunnels at NASA Langley Research Center and second, to apply them in the Advanced Propulsion Laboratory (APL) (supported by a AFOSR Grant) at Hampton University. The developed FS system is already being applied in the HFG facility at NASA LaRC to visualize low-density flow in the free-jet regime.

TECHNICAL DISCUSSION

A. Electron-Beam Fluorescence

A schematic of our EBF system is shown in Figure 1. The system mainly consists of a commercially available electron gun, a stainless steel vacuum chamber, vacuum pumping system, spectroscopic system, gas supply, and focusing optics. One or two gases can be mixed and flowed into the continuously-pumped vacuum chamber until a static pressure is reached. The electron gun, maintained at a pressure of less than 104 Torr, can then be energized between 100 eV and 10 keV. The beam of electrons (between 1 and 2 mm diameter) traverses the chamber, passes through the gas at the center of the chamber, and is collected by a faraday cup, which collects the beam current. The fluorescence emission produced is then observed through an 80 mm diameter quartz window and focused, with a 150 mm quartz lens, into the entrance slit of a 0.22 m double spectrometer. The 0.22-m double spectrometer provides dispersion (with two 1200 gr/mm holographic gratings) at the exit slit where a photomultiplier tube is mounted to amplify the signal. The spectral information can then be sent to a computer where it was displayed and analyzed

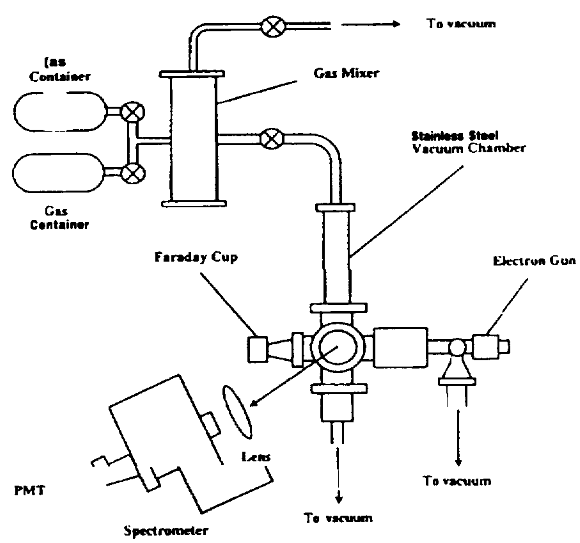


Figure 1. Diagram of the Electron Beam Fluorescence system.

The EBF system was calibrated with N_2 gas to evaluate its sensitivity and reliability for application to high-speed aerodynamic flowfields. Among many channels of electron-N_2 collisional kinetics, excited molecular nitrogen ions, N_2^{+*}, are mainly produced in the reaction

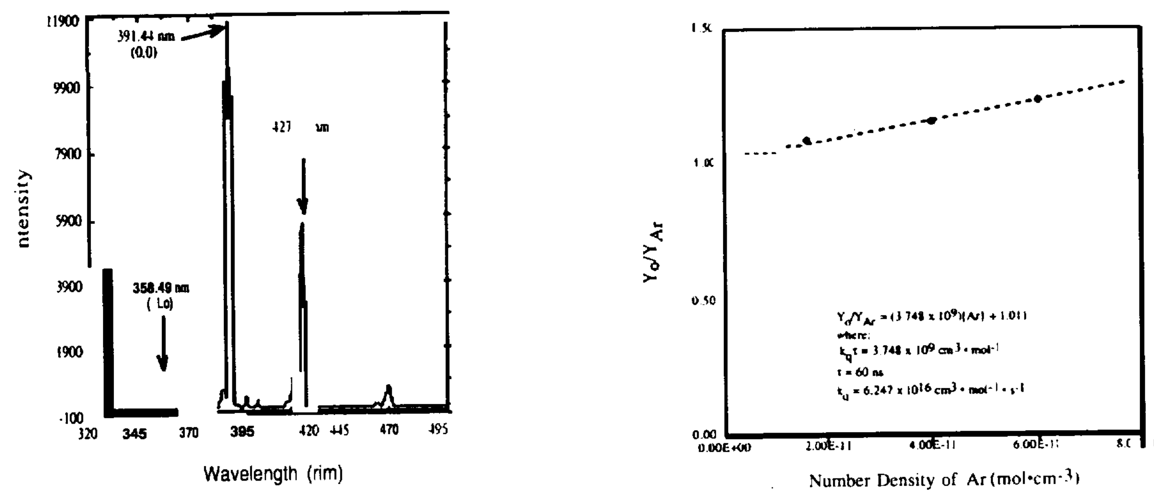
$$e" + N_2 \dashrightarrow N_2^{+*}(B^2\Sigma_u^+) + 2e^-.$$

This is followed by the emission of fluorescence in the First Negative band,

$$N_2^{+*} \dashrightarrow N_2^+(X^2\Sigma_g^+) + h\nu$$

with the (0,0) vibrational transition (at λ = 391.44 nm) being the most intense feature and considered to be a sensitive probe for N_2 concentration.

Figure 2 shows some of our preliminary experimental results. Figure 2a is a typical emission spectrum of N_2^+ emission in the First Negative band between 320 and 500 nm. We were also able to determine the quenching rate of N_2^+ emission (at λ = 391.44 nm) by argon in order to shed some light **on** the radiative properties of hypersonic aerospace vehicle bow-shock layers (see Figure 2b). We calculated a quenching rate, k , in the order of 10^{16} $cm^3mol^{-1}s^{-1}$. This is two orders of magnitude higher than we anticipated (k_q is typically less than 4.5 x 10'" cm^3 mol^{-1} s"' for N_2^+ + Ar). The source of our error is being investigated.



(a) (b)

Figure 2. (a) Typical fluorescence emission spectrum of N_2^+ First Negative Band. (b) Curve of N_2^+ emission (at λ = 391.44 nm) quenching by argon.

B. Focusing Schlieren

The Focusing Schlieren technique has several advantages for flow visualization because it produces a natural, easily-interpretable image of refractive-index-gradient fields. This low-cost technique also has focusing capability, and high sensitivity. Figure 3 shows a layout of the optical components of a small-field, high-sensitivity Focusing Schlieren system. This system consists of a source grid and a cut-off grid (located on opposite sides of the flowfield), a light source, a Fresnel lens, an image lens, and an intensified CCD camera. Refraction of the collimated light beam in the flow region moves the image of the source relative to the knife edge, which results in a change in the brightness at the image of the test section.

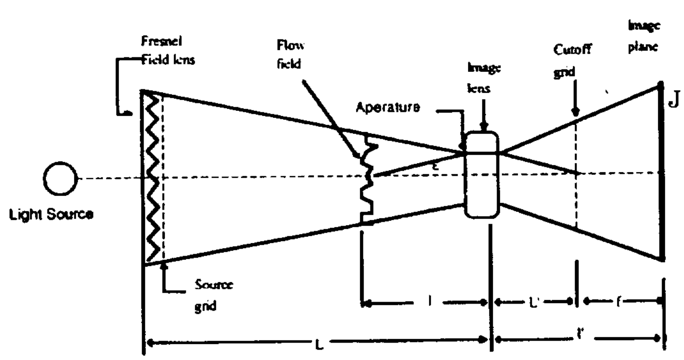


Figure 3. Layout of optical components of a small-field, high-sensitivity Focusing Schlieren system.

A Focusing Schlieren system is characterized by its sensitivity, resolution, and depth of focus[7]. The sensitivity of the system is defined by the angle of deflection normal to the knife edge. If this angle is shown by ε', the image of the source shifts by $\Delta a \approx \varepsilon' L'$, where a is the light source image height above the cut-off grid, and L' is the distance from the lens to the cut-off grid. This results in a change in intensity

$$\frac{\Delta I}{I} = \frac{\Delta a}{a} = \varepsilon'\left(\frac{L'}{a}\right)$$

If we use the criterion that the smallest change in brightness that can be detected is 10%, the sensitivity of the system becomes $\varepsilon' \approx 0.1\ (a/L')$, and this quantity is defined as

$$\varepsilon'_{min} = 20626\ \left(\frac{a}{L}\right)\ \text{arcsec}$$

Due to the non-parallel nature of the light, the sensitivity of the Focusing Schlieren is:

$$\varepsilon'_{min} = \frac{20626aL}{L'(L-l)}\ \text{arcsec}$$

where, L is the distance from the source grid to the lens and l is the distance from the flowfield object to the lens.

The following parameters specify the sensitivity of our system:
$L = 7.6$ m, $L' = 1.17$ m, $l \approx 4.6$ m, and $a = 0.038$ mm. Sensitivity $\varepsilon'_{min} = 1.7$ arcsec.

Figure 4 shows a diagram of the small-field, high brightness, and highly-sensitive Focusing Schlieren system incorporated into the High-Speed Flow Generator (HFG) at NASA LaRC. This system is able to visualize and analyze low density high-speed flows. Of particular interest is the visualization of the boundary layers associated with a continuum free jet expansion.

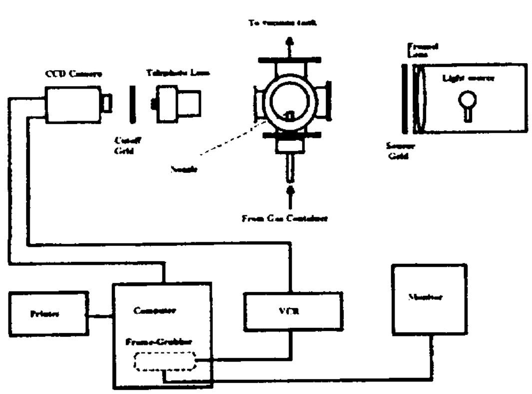


Figure 4. Diagram of the highly-sensitive Focusing Schlieren system incorporated into the High-Speed Flow Generator (HFG) at NASA LaRC.

Figures 5a and 5b show preliminary results of the application of this system to low density

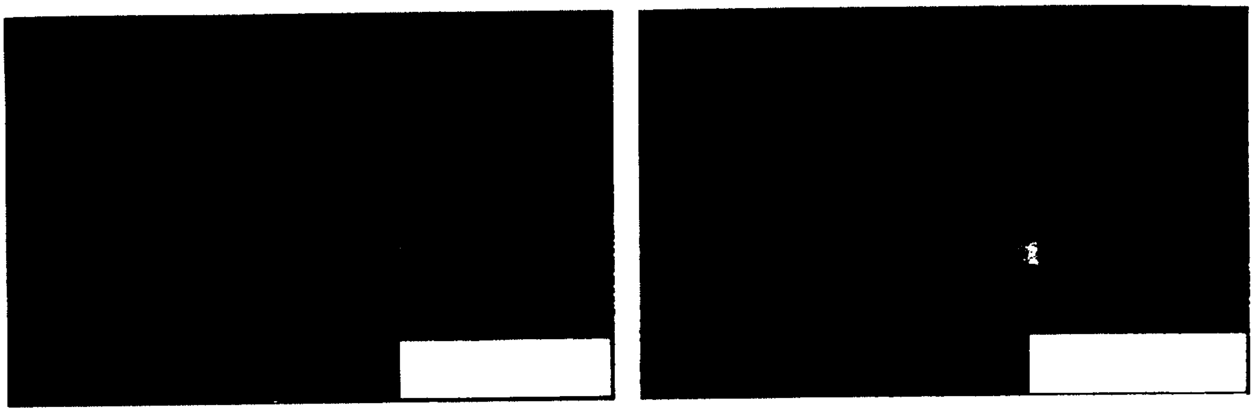


Figure 5 (a) Background-subtracted VCR picture of the HFG system with a 5.0 mm nozzle. Free-jet expansion of argon gas (at 200 psig) into a 22 micron vacuum is visualized with a highly-sensitive Focusing Schlieren system. (b) Back-ground subtracted VCR picture of the HFG system with 5.0 mm nozzle and a spherical barrier. Argon gas is flowed (at 200 psig) into a 22 micron vacuum and shock-front layers around the sphere are visualized with the Focusing Schlieren system.

C. Laser-Induced Fluorescence

Figure 6 shows the layout of LIF diagnostic of the APS, namely, the Solar Thermal-Electric Propulsion (STEP) system. A tunable dye laser (Lambda Physik, LPD 3002CES: Tuning range 332 nm - 860 nm) pumped by an excimer laser (Lambda Physik, LPX 220I CC: Computer controlled, maximum energy of 200 mJ, pulse duration of 12 ns, 60 Hz) is in place and has been tested for operation. The exhaust flow speed from the STEP is expected to reach several km per sec., providing an ideal testbed for the LIF technique. The dye laser system has produced an output of 11 mJ per pulse at 570 nm when Rhodamine 6G was used in a preliminary test run.

In order to measure the density, velocity, and temperature of the high-speed exhaust flow in the STEP system, we will use the intensity, Doppler shift, and broadening of the 728.7 nm fluorescence line of argon gas, respectively. This task is under way and the expected results will be reported later.

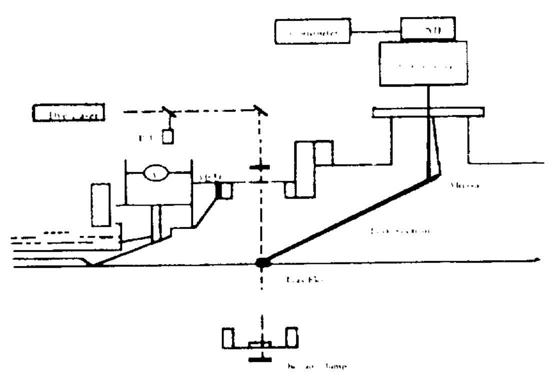


Figure 6. Layout of LIF Diagnostic of STEP system

ACKNOWLEDGMENT

The authors would like to acknowledge Dr. Richard Antcliff, Head of the Measurement Science and Technology Branch, NASA-LaRC and the staff of the Branch for their continuous support of our work. Also, we would like to acknowledge Dr. Doyle Temple, Chairman of the Department of Physics and Director of RCOP for his interest and support of our research.

REFERENCES

Charles A. Terrell, Bagher M. Tabibi, and Ja H. Lee ", Bulletin of the APS, November 14-16, 1996, Vol. 41, No. 8, AC 2.

2. Muntz, E. P., "The Electron Beam Fluorescence Technique," AGARDograph 132, 1965; Honaker, W. C. et all, "A Study of Density Measurements in Hypersonic He Tunnels Using an EBF Technique," AIAA 14th Thermophysics Conf., June 4-6, 1979, Orlando, Florida ; Cattolica, R. J., et al , Physics Fluids 17, 1713, 1979.

3. Hiller, B. et al, "LIF Technique for Velocity Field Measurements in Subsonic Gas Flows," Optics Letters, Vol. 8, No. 9, 1983, pp. 474-476; McKenzie, R. L., et all. Applied Optics, Vol. 20, No. 12, 1981, pp. 2153-21-65. see also Optics Letters, Vol. 8, No. 7, 1983, pp. 368-370; Miles, R. B., et all, Optics Letters, Vol. 13, No. 3, 1988, pp. 195-197; Cohen, L. M., et al, AIAA Paper 87-1527, June 1987.

4. Williams, W. D., et al, "Raman and Rayleigh Scattering Diagnostics of Two-Phase Hypersonic N_2 Flowfield," AIAA Journal, Vol. 13, No. 6, 1975, pp. 709-710 ; Shirizade B., et al, "The Impact of Condensation Effects on Rayleigh Scattering in a Mach 6 Wind Tunnel," NASP TM 1102, May 1990; Miles, R. B., and Lempert, W. "Two-Dimensional Measurement of Density, Velocity, and Temperature in Turbulent High-Speed Air Flows by UV Rayleigh Scattering," Applied Physics B, Vol. B51, No. 1, 1990, pp. 1-7.

5. Lederman, S., "The Use of Laser Raman Diagnostics in Flow Fields and Combustion," Journal of Progress in Energy and Combustion Science," Vol. 3, No. I-A, 1977, pp. 1 -34; Exton, R. J., et al, "Molecular Flow Velocity Using Doppler Shift Raman Spectroscopy," AIAA Paper 87-1531, June 1987; Antcliff, R., et al "A Hardened CARS System Utilized for Temperature Measurements in a Supersonic Combustor," AIAA Paper 91-0457, Jan. 1991; Cattolica, R. J., et al, AIAA Paper 90-0627, Jan 1990.

6. Gartrell, L. R., et al, Laser Velocimetry Technique Applied to the Langley 0.3-Meter Transonic Cryogenic Tunnel," NASA TM-81913, 1981

7. Weinstein, L. M., "An Improved Large-Field Focusing Schlieren System," AIAA Paper 91-0567, Jan. 1991; see also "Large-Field High-Brightness Focusing Schlieren System," AIAA Journal, Vol.31, No. 7, 1993, pp. 1250-1255; AIAA Journal, Vol. 32, No. 6, 1994, pp. 1242-1249.

5/23/33

URC97123

Single Event Upset Immune CMOS SRAM By Circuit Design

Tian-Shen Tang and John S. Linder
Department of Electrical Engineering and Computer Science
Texas A&M University-Kingsville, Kingsville, Texas 78363

Abstract This paper discusses some issues related to designing a single event upset (SEU) tolerant CMOS SRAM using a commercial VLSI process. The techniques used for the design are on-chip error correction and active resistance SEU hardening. The key difference of this approach from the traditional techniques is that it achieves SEU protection via a circuit design approach while the traditional approaches are mainly based on radiation tolerant processes. Since no process modification is required, a commercial IC process can be used to develop SEU immune products when the total dose effects are not a concern.

I. Introduction

This paper discusses some issues related to designing an SEU tolerant CMOS SRAM using a commercial VLSI process. The techniques used for the design are on-chip error correction and active resistance SEU hardening [1,2]. The on-chip error correction technique employs a linear block coding (LBC) circuit to correct single-bit errors in the CMOS SRAM. This technique, capable of detecting double-bits errors, corrects errors by flipping the error bit rather than rewriting the entire corrected code word which reduces the writing time and the complexity of the circuitry. It also eliminates the problems caused by error accumulation. Since the LBC circuit can only safeguard the data stored in the SW, the data which are used by the LBC circuit has to be protected from SEU effects by other means. The data which can be generated, such as the addresses of the bits of a code word, are generated instead of being stored, so that the duration of SEU upsets will be very short. The data which has to be stored are stored in SEU immune buffers. The SEU immunity of these buffers is provided by transmission gate (TG) like active SEU hardening resistors.

Error control coding techniques have been used for correcting SEU induced errors in SRAM memory. The information stored in the memory is organized as code words which contain data bits and check bits. Error detection and correction logic generates the check bits upon a write operation and uses the data and check bits upon a read operation for error detection and correction. There are many drawbacks in such an approach. A data word is fed to the error detection and correction circuit for generating new check bits only when a bit of that word is used. Therefore, if that word is not used for a long time, errors could be accumulated in the code word and become uncorrectable. Moreover, it takes a long time before a bit becomes available in a read operation because the error detection and correction must be performed first. After errors have been corrected, the entire code word is written back to the memory, which is time consuming. The proposed approach will not have these problems.

This paper is organized as follows. The organization of the SEU immune CMOS SRAM and the architecture of error detection and correction circuitry are described in the second section. The aspects of the major components of the SRAM are described in the third section. Active resistor hardening technique is discussed in the fourth section. The layout of an SEU immune memory cell is given in this section. Conclusions are given in the last section.

II. Organization of the SEU Immune SRAM

A prototyping SEU immune SRAM of the proposed architecture is being designed. It consists of the following functional blocks (Fig. 1): the storage array, the address generation the row and column decoders, the read/write circuit, the parity generation, the error detection and correction and the input data control.

1. *The storage array is* divided into two 16× 16 SRAMs, one for data bits and one for check bits. An (8,4) LBC code, consisting of 4 data bits and 4 check bits, is used. The check bit SRAM is not accessible by the user. To the user, the chip is just a conventional 16× 16 SRAM. Each SRAM is divided into 4 different

blocks. Different bits of the code words are stored in different blocks in their respective positions. A detailed discussion is given in the next section.

2. *The address generator* produces the addresses of all the bits in a code word according to the address of a reference bit. The addresses of all the reference bits are generated by a sequential ring counter. The check bit SRAM is in parallel with the data SRAM. Only one address generation is sufficient. Another function of the generator is to determine the address of the reference bit if one of the data bits of its code word is to be written.

3. *The row and column decoders* are the same as those of conventional SRAMs, decoding the row and column addresses of the bit to be read or written.

4. *The read/write circuit* has the same functions as those of conventional SRAMs, performing read/write operations. No error detection or correction is performed at this stage, so that the data can be quickly accessed from or sent to the data SRAM.

5. *The parity generation* generates the parity check bits of the code words. When a bit in the data SRAM is written by the user, a new code word has to be generated and stored. According to the address of the reference bit of this code word provided by the address generator, the parity generator works with the read/write circuit to load the remaining three data bits of the code word and generates the four parity check bits, and then stores the new parity bits.

6. *The error detection and correction* is continuously working under the supervision of the input data control circuit to detect and correct single-bit errors that occurr in both the data and parity check SRAMs.

7. *The input data control sends an* interrupt signal to pause the detection and correction process and switches the control to the read/write circuit, when there is a Read or Write request from the user. Once a read or write operation completes, the control circuit checks for any another request before it continues with the error detection and correction process.

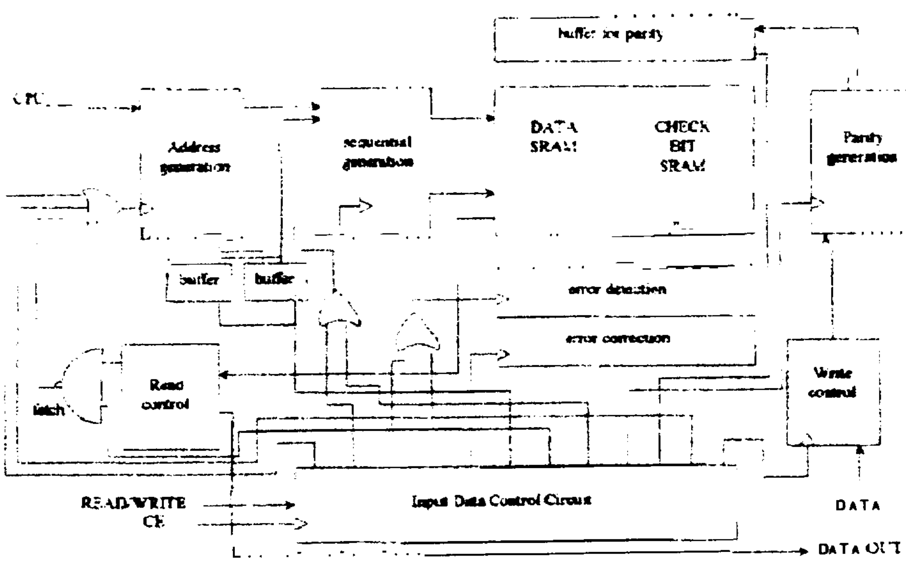


Figure 1. Organization of the SEU immune SRAM with on-chip error detection/correction

III. Special Aspects of Some Major Functional Blocks

3.1 Error Detection and Correction

The used error correcting code is an (8,4) linear block code [3]. The advantage of using LBC is its linear systematic structure which reduces the encoding complexity. A code word is divided into two blocks: the data bits block and the parity bits block. The stored data can be read directly from the data SRAM without decoding. When a bit is written a new code word needs to be generated. In general, if data to be encoded is $\mathbf{u} = (u_0, \ldots\ldots u_{k-1})$ then the code word v is given by $v = \mathbf{u}.\mathbf{G}$ where the binary matrix G is called the

generator matrix. Since the matrix G is of the form $\mathbf{G=[I_k\ P]}$, encoder complexity is greatly reduced. The G matrix of the LBC used in this paper is given below.

$$\mathbf{G}=\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Another important matrix of the LBC is the 'parity check matrix' **H** which is of the form $\mathbf{H}=[P'\ \mathbf{I}_{n-k}]$. The product vector $s = \mathbf{d}.\mathbf{H}^T$ **is** called the syndrome, where $\mathbf{d=v+e}$. **in** which v is a code word and e is an error vector. For any code word v, $\mathbf{v}.\ \mathbf{H}^T\mathbf{=0}$. Therefore, s = O indicates no single and double errors; otherwise an error is detected, Let us denote the *i*th bit error vector by $\mathbf{e}_i$ whose *i*th bit is one and others are zero. A single error at the *i*th bit of a code word v can be described as $\mathbf{d=v+e}_i$, and its corresponding syndrome is the *i*th row of matrix $\mathbf{H}^T$ **since** $\mathbf{s=(v+e_i).H}^T=\mathbf{e_i.H}^T$. Assume that only single-bit errors occur and the number of parity bits n-k satisfies $2^{n-k}\geq n$. A demultiplexer which uses syndrome as control inputs can be used to locate the position of a single error. Once the error bit has been detected, it can be corrected easily by flipping the error bit since $\mathbf{e_i+e_i=0}$ and $\mathbf{d+e_i=v+e_i+e_i=v}$. As a result, only one bit of the corrected code word needs to be written to the cell, rather than the whole corrected code word. Figure 2 shows the schematic of the single error detection and correction circuit, The main components used are data buffers, EXOR gates, and demultiplexers.

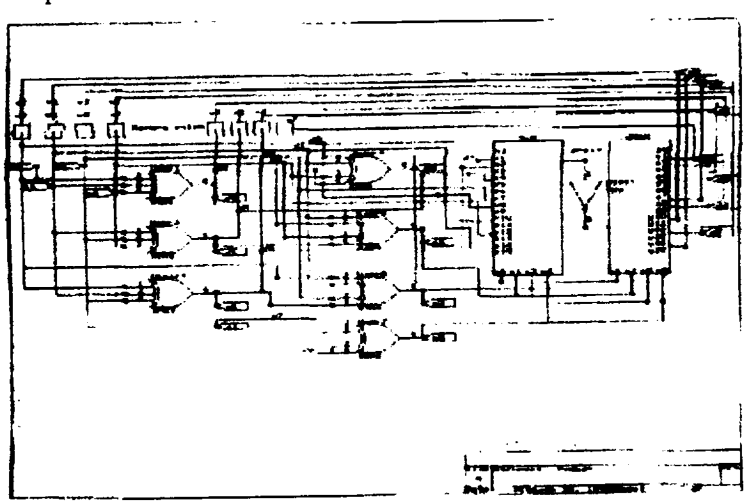


Figure 2. Error detection and correction circuit

3.2 **Arrangement of the SRAMs**

The occurrence of multiple errors in a single code word due to a single event upset must be made less probable, since the circuit can correct only single errors. The occurrence of multiple errors can be minimized by appropriate storage arrangement of the code words in the memory [4]. The storage arrangement used for this purpose is shown in Fig. 3. This kind of arrangement takes the advantage of the regular matrix structure of the SRAMs. The bits of the code word are physically separated and are placed in different addressing locations following a regular pattern. Each SRAM is logically divided into four quadrants where for this purpose a 16× 16 SRAM is separated into four quadrants of 4×4 each. For this consideration a code word *consists* of 8 *bits* and each bit is *stored in* the corresponding locations in each quadrant of the two SRAMs.

Of the two SRAMs, only the data bit SRAM can be accessed by the user. The parity check bit SRAM can only be accessed by the error detecting and correcting circuitry and cannot be accessed by the user, thus reducing the circuit complexity. A special address generation logic is necessary in order to access the bits.

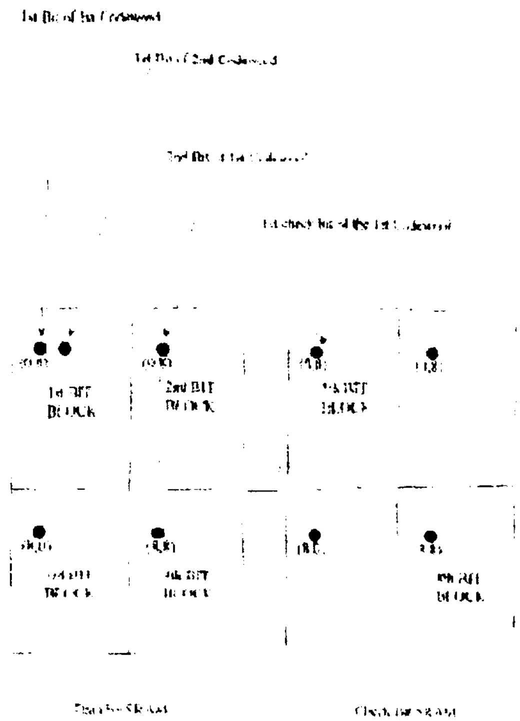


Figure 3 Arrangement of the SRAMs

3.3 **Address Generations**

The address generation circuit has two major tasks: generating the bit addresses of all the code words for error detection and correction and generating bit addresses for encoding.

ADDRESS GENERA TOR FOR ERROR CORRECTION The error correction circuit scans the stored code words sequentially. In order to read the stored code words and to write back the corrected code words, the absolute addresses of the code words in the memory must be generated. First and second bits of a code word are in the same row but separated by 7 intervening columns. Similar relation holds for third and fourth bits. Therefore with the combination of addresses of only two bits (first and fourth), we can generate addresses of other two bits. Besides, the check bits of the same code word were stored at the same addresses in the second SRAM; therefore, only two address buffers are needed which simplifies the address generation logic. Scanning the stored code words is carried sequentially starting with the first bit of each code word as the reference. After scanning all the code words, the control goes to the initial address. Thus, this technique has *the* capability to check each bit error in *close range.*

ADDRESS GENERA TION FOR ENCODING A write cell signal initiates an encoding operation and causes the current error correction operation to hold. Like error-correction operation, the absolute addresses of the four associated data bits in the data SRAM must be generated. Unlike the error correction operation, the written cell can be any bit of the four so that the addresses of the first and fourth bits are generated and stored. The generation of the frost address is based on the address of the written cell (x, y) and the block address formed by the most significant bits of the binary forms of x and y i.e., (x_3, y_3). If $(x_3, y_3)=(0, 1)$, it indicates the written cell is the second bit, then the absolute address of the first bit is (x, y-8). That is the same row and the eighth column to the left of the written bit. The fourth bit address is $(x+8, y)$. With these addresses the four data bits in the code word is read and stored in data buffers and later given to the parity generator to generate check bits. The four generated check bits are stored to the check bit SRAM with the same addresses of the data bits. For this operation four data and two address buffers are sufficient.

4. The Active Resistor SEU Hardening Technique

As seen in the last section, data and address buffers are used to stored the generated addresses and data bits

for encoding and error correction operations. To protect the data stored in the buffers from SEU upsets, the active-resistor SEU hardening technique is used. A SEU immune buffer cell is acturally is a SEU immune SRAM ceil with transmission gates (TG) as feedback resistors connected between the two CMOS inverters, as seen in Fig. 4.

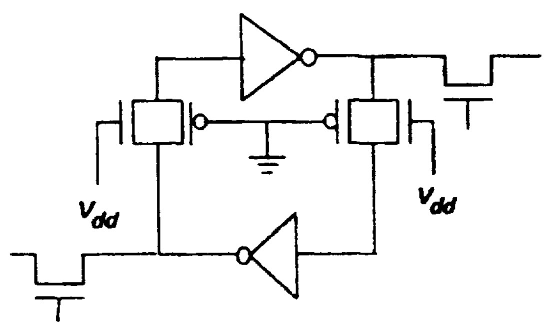


Figure 4 A CMOS SRAM Cell with TGs as Feedback Resistors

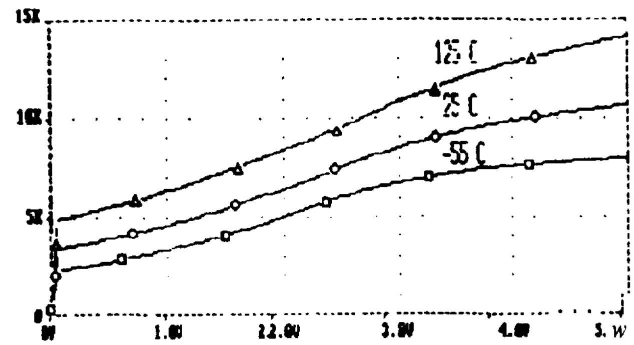


Figure 5 TG Resistance vs. Terminal Voltage and Temperature (L=2μ, W=2μ)

A CMOS TG exhibits a nonlinear current-voltage characteristic when it conducts; hence called nonlinear active resistor. Figure 5 depicts resistance as a function of terminal voltage for a CMOS TG that consists of two minimum size MOSFETs (MOSIS 1.2pm design rules). As seen in Figure 5, the TG resistance strongly depends on its terminal voltage increasing rapidly as the terminal voltage increases. This phenomenon can be utilized to increase SEU immunity of a SRAM cell. A SEU hardened CMOS SRAM cell using TGs as feedback resistors is shown in Figure 4. In this cell the inverter pair is decoupled by two TGs whose p-channel and n-channel transistors are respectively gated by the ground and power source. These transistors provide the resistance needed for increasing critical charge of the cell and also introduce additional capacitance to the sensitive nodes and feedback paths of the cell, which can effectively increase SEU immunity of the cell, as shown by us and others [5,6].

Operation of the cell can be described briefly as following. When the cell operates normally, the resistance of the two TGs is very low since the voltages across the gate terminals are very small. The cell is essentially an unhardened one. When one of the sensitive nodes is hit by an ionizing particle, electrical charges are collected at the hit node, causing a sudden voltage increase or decrease at the hit node while the voltages at other nodes are relatively unaffected. In response to the voltage increase across the terminals of the TG connected to the hit node, the resistance of the TG becomes very high. The high feedback resistance protects the stored cell data from SEU. Figure 4 shows a basic configuration, other alternatives which provide more SEU upset protections and fast speed performance were developed from this basic configuration [7]. A layout design of such a SEU immune SRAM cell is shown in Fig. 6.

5. Conclusions

Some issues related to designing an SEU tolerant CMOS SRAM using a commercial VLSI process have been discussed. This approach employs the on-chip error correction and active resistor SEU hardening techniques to **achieve** SEU immunity and is a circuit design approach. The error correction circuit to correct the single errors caused by SEUs is shown to be simple and therefore, the writing time is reduced. Address generation logic is simplified when we divide the memory into bit blocks. The incorporation of the error correct ion and active resistor SEU hardening on-chip does not require any changes in the fabrication process and in the way that data is stored and read. A commercial process can be used to fabricate the chip. This SRAM can be used in the enviornments where total dose effects are not serious problems. For example, a short space mission which lasts for only a few weeks. The detailed design of the circuit is under way. The chip will be fabricated and tested.

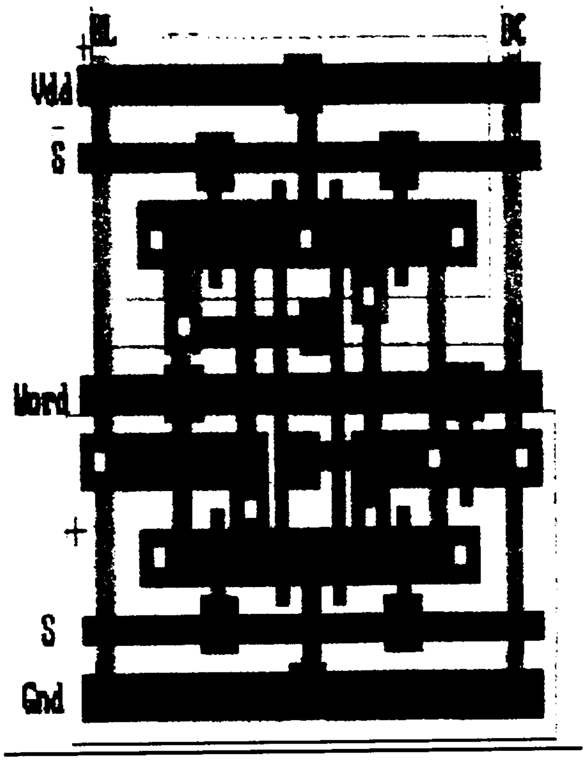


Figure 6. Layout Design of A CMOS **SRAM** Cell with **TGs** as **Feedback** Resistors

REFERENCES

1. T.S. Tang, I.P. Tadpati, and J.S. Linder, "An active resistor-hardening technique for CMOS SRAMs," Engineering& Architecture Symposium'93, Prairie View, TX, March 15-16, 1993.
2. G. Merugumuvvala, *Architectural Design of SRAM with On-Chip Error Detection and Correction Against Single Event Upset,* MS Thesis, Texas A&M University-Kingsville, 1995.
3. Shu Lin and Daniel J. Costello, Jr., *Error Control Coding: Fundamentals and Applications,* Prentice-Hall, 1983.
4. Zoutendyk, H.R. Schwartz, R.K. Watson, Z. Hasnain and L.R. Nevill, "Single Event Upset (SEU) in a DRAM with on-chip error correction", *IEEE Trans. on Nuclear Science,* Vol NS-34, No.6, pp. 1310-1316, Dec 1987.
5. Tang, S. Yeluru, and J.S. Linder, "Single event upset immune CMOS SRAMs with distributed RC feedback," Engineering & Architecture Symposium' 93, Prairie View, TX, March 15-16, 1993.
6. T. Iizuka and T. Sakurai, " CR isolated cell for soft error prevention-static RAM application," *IEEE 83 Symposium on VLSI Technology.*
7. S.M. Patamalla, *Layout Design of SEU Immune CMOS Static RAM Cells with Active Feedback Technique,* MS Thesis, Texas A&M University-Kingsville, 1995.

URC97124

Au Colloids Formed by Ion Implantation in Muscovite Mica Studied by Vibrational and Electronic Spectroscopes and Atomic Force Microscopy

Y. S. Tung[a], D. O. Henderson[a], R. Mu', A. Ueda[a], W. E. Collins, C. W. White[b], R. A. Zuhr[b], and Jane G. Zhu[b]

'Chemical Physics Laboratory, Physics Department, and NASA Center for Photonic Materials and Devices, Fisk University, Nashville, TN 37208
[b]Oak Ridge National Laboratory, P. O. Box 2009, Oak Ridge, TN 37831

Abstract

Au was implanted into the (001) surface of Muscovite mica at an energy of 1.1 MeV and at doses of 1, 3, 6, and 10x 10^{16} ions/cm^2. Optical spectra of the as-implanted samples revealed a peak at 2.28 eV (545 nm) which is attributed to the surface plasmon absorption of Au colloids. The infrared reflectance measurements show a decreasing reflectivity with increasing ion dose in the Si-O stretching region (900-1200 cm-l). A new peak observed at 967 cm^{-1} increases with the ion dose and is assigned to an Si-O dangling bond. Atomic force microscopy images of freshly cleaved samples implanted with 6 and 10x10^{16} ions/cm^2 indicated metal colloids with diameters between 0.9- 1.5 nm. AFM images of the annealed samples showed irregularly shaped structures with a topology that results from the fusion of smaller colloids.

1. Introduction

The reduction of matter to the nanometer scale often leads to a manifestation of properties which differ significantly from the bulk. The thermodynamic, linear and nonlinear optical, structural, and mechanical are examples of properties which are modified when a materials is reduced to a finite dimension. Ultrafine scale gold particles have drawn much interest due to appearance of the surface plasmon polarition which is observed when the particle is reduced to the nanometer scale. [1] More recently gold nano-particles have shown to possess a large nonlinear optical response and have properties that are desirable for an all optical switching device .[2]

Several approaches have been taken to fabricate metallic nano-particles which include hydrosols, impregnation/chemical reduction, inert gas evaporation, vacuum evaporation, vacuum evaporation and matrix isolation, cluster beams, reverse micelles, and pressure impregnation.[3] We have taken the approach of using ion implantation into the (001) surface of mica to form colloids. The use of mica as a substrate has the advantage that it is atomically flat which allows for AFM measurements to be carried out to determine the colloid size in the as-implanted state and to follow the effects of annealing on colloid growth and aggregation. In addition, mica is reasonably transparent in the visible region and therefore, as a substrate, its permits the study of the surface plasmon resonance of Au nano-particles.

2. Experimental

Au implantation into (001) oriented mica at room temperature was carried out at energies of 1.1 MeV and at doses of 1, 3, 6, 10x 10^{16} ions/cm^2. Thermal annealing was carried out at

temperatures between 200 and 5000 C in a reducing atmosphere (5% H_2 + Ar) with a one zone tube furnace.

The electronic spectra were measured between6.2-1.24 eV (200-1000 nm) with a Hitachi model 3501 UV-Vis-NIR spectrophotometer. A piece of virgin mica of the same thickness as the implanted sample was placed in the reference beam for normalization of the spectra. The resolution of the measurements near 550 nm was 0.2 nm. Repeated measurements showed the photometric accuracy to be ±0.3 % transmittance and ±0.2 nm in wavelength accuracy. The band centers were located from the extrema in second order derivative spectra.

The infrared reflectance spectra were obtained with aBomem MB- 100 Fourier transform spectrometer to cover the 4000-400 cm-1 range. The samples were mounted on a laboratory built reflectance stage with a fixed angle of incidence of 150 .Typically, 200 interferograms were collected. The wavelength and photometric accuracy were 0.5 cm^{-1} and 0.4 % reflectance, respectively. All spectra were divided by a reference spectrum of a gold mirror. The band centers were obtained from extrema in the second derivative spectra.

All AFM images were recorded in the tapping mode with a Nanoscope III atomic force microscope from Digital Instruments. The height resolution reported for all measurements is ± 0.1 nm. Samples were prepared for AFM measurements by applying adhesive tape to the implanted surface and stripping away layers of mica. Both the mica adhering to the tape and that which remained on the implanted substrate were examined by electronic spectroscopy to identify where the colloids were located by the appearance of the surface plasmon polariton.

3. Results

The infrared reflectance spectra for mica implanted with Au at doses of 1, 3, 6, and 10×10^{16} ions/cm^2 are shown in figure 1 together with the reflectance spectrum of virgin mica. The spectrum for virgin mica shows two intense reflectance maxima located at 1096 and 1041 cm^{-1} and an unresolved shoulder at 1014 cm^{-1} and a weak peak at 904 cm^{-1}. The Au-implanted substrates show a decrease in reflectivity with an increase in ion dose. The corresponding bands in the ion implanted samples are redshifted from those of the virgin mica; the 1096 cm^{-1} peak shifts by 21 cm^{-1} to 1075 cm^{-1} and, the 1041 cm^{-1} peak shifts by 16 cm^{-1} to 1.025 cm^{-1} for the implanted samples. There is an appearance of a new peak at 967 cm^{-1} that is initially poorly resolved for the lowest ion dose, but it becomes more intense and better resolved as the ion

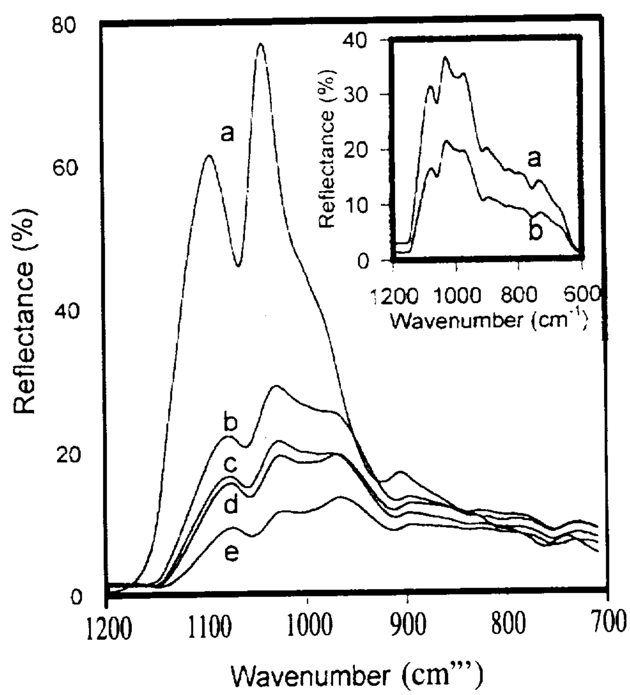


Figure 1. Infrared reflectance spectra for Au implanted: a) virgin, b)1e16, c)3e16, d)6e16 and 10e16 ions/cm^2. Inset 6e16 as-implanted and annealed at 500°C in 5 % H_2 +95% Ar.

dose increases. The 1014 cm-l feature observed in the virgin mica cannot be observed in the implanted samples because it is masked by the dominant intensity of the 967 cm^{-1} band.

The affect of annealing on the infrared spectra of the implanted samples is illustrated in the inset of figure 1. There is a significant increase in the reflectivity of for the annealed sample but, the peak at 967 cm^{-1} remains.

The electronic spectra are shown figure 2 for the mica substrates implanted with 1, 3, 6 and $10x10^6$ ions/cm^2. The prominent feature in the spectra appears at 2.28 eV (545 nm) and increases with the ion dose, There is a dose dependent increase in absorption at energies greater than 3 eV and virtually no absorption is observed for energies less than 1.2 eV. Annealing the samples at 5000 C results in an increased absorption at 2.28 eV and this is shown for the sample implanted $6x10^{16}$ ions/cm^2 in the inset of figure 2.

An AFM image of a freshly cleaved unannealed mica surface implanted with $10x10^{16}$ ions is shown in figure 3. The areas of high contrast appearing as white spikes correspond to heights between 1.03-0.59 nm and are attributed to Au colloids. Further analysis of the data indicates that 86% of the colloids are 0.88 nm in diameter, while 8% are 0.59 nm and the remaining 6% are 1.03 nm in diameter.

The results of a sample annealed at 500°C for 1 hour are shown in figure 3. The main feature observed in the image is the area of white contrast which shows an irregular shaped feature -1 nm in height, The image shows domains that suggest that irregular shaped feature is composed of smaller clusters that have begun to fuse together during the annealing.

Figure 4 shows an image of an annealed sample implanted with $10x 10^{17}$ ions/cm^2 where two (001) planes are separated by several steps (40 rim). It is clear from the image that colloids on the two planes differ significantly in size.

4. **Discussion**

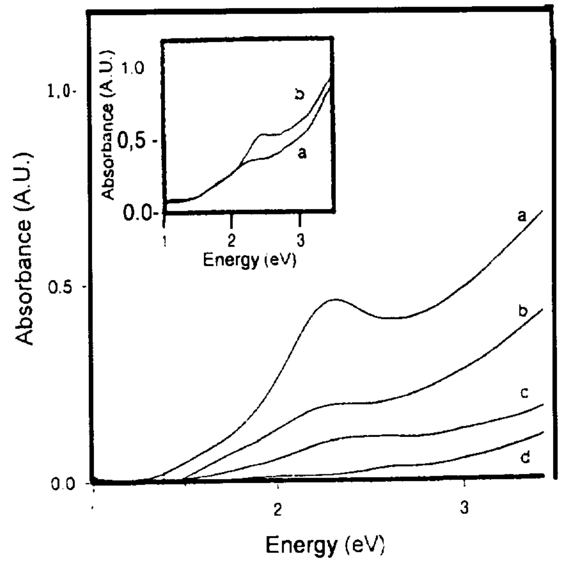


Figure 2. Electronic spectra for Au implanted in mica at 10e16, 6e16, 3e16 and 1e16 ions/cm^2 from top to bottom. Inset shows the effect of annealing the same sample under the same conditions as shown the inset of figure 1.

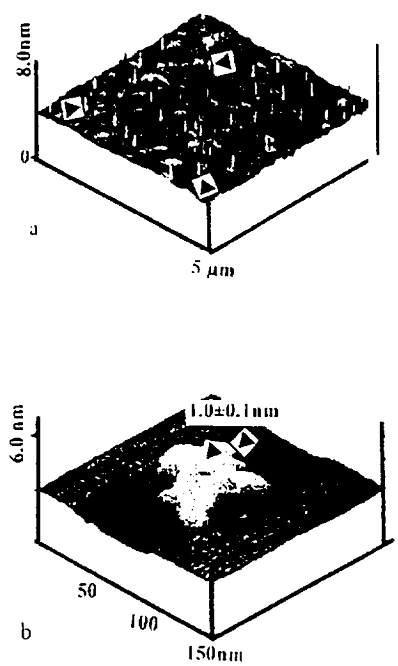


Fig 3. a) An AFM image of Au implanted in mica (10e17). The same sample annealed at 500° in 5 % H_2 + 95% Ar. Arrows indicate height of the cluster above the (001) plane.

The peaks observed for the virgin mica in the infrared spectra between 900-1200 cm-1 can be attributed to Si-O stretching vibrations.[4] The decrease the frequencies and intensities of the Si-O stretching vibrations which occurs with the concomitant appearance of the peak at 967 cm * is attributed to a decrease in the number of oscillators that absorb in 1100- 1200 cm" [1] region. The decrease in the number of Si-O oscillators is in turn attributed to the ion induced rupture of the Si-O-Si bridge bonds. The red shift in the Si-O stretching vibrations is due to a decrease in the force constant associated with the Si-O stretching vibration and also to changes in the G-matrix elements in the Wilson secular equation. [5] The peak observed at 967 cm-l in the Au-implanted mica sample correlates reasonably well with the frequency assigned to the Si-O dangling bond vibration for fused silica implanted with heavy ions.[6] On this basis we attribute the peak at 967 cm^{-1} in Au implanted mica to a Si-O dangling bond defect. Annealing the samples restores some of the intensity to the Si-O stretching region, but the persistence of the peak at 967 cm^{-1} indicates that the extensive defects remain in the lattice.

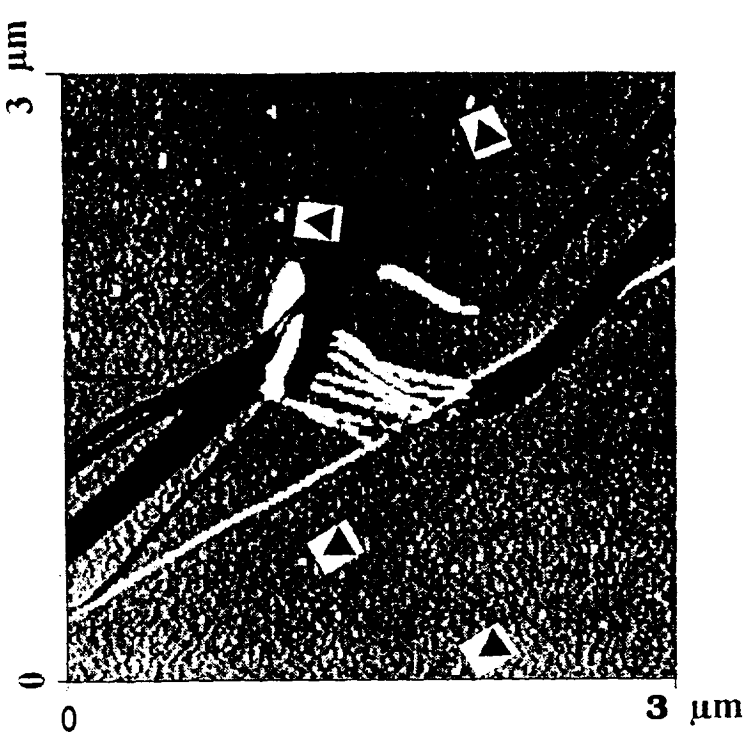


Fig 4. An AFM imageof mica implanted with 1e17 ions/cm^2. The arrows point to colloids to colloids separated on different planes that are separated by 40 nm.

The most significant feature in the electronic spectra is the peak at 2.28 eV which grow on annealing. By comparison to other studies Au implanted into sapphire and fused silica where the surface plasmon polarition is observed at 2.25 eV,[7] we assign the 2.28 peak in mica to the surface plasmon polariton to Au colloid in mica.

Figure 4 shows the Au colloids on two (001) planes of mica separated by 40 nm. The data are presented with z-axis in terms of deflection for clarity. It is clear from the image that the colloids on the upper plane are much larger than those in the lower plane. The difference in the sizes of the colloids must be due to that fact that the implanted profile has a Gaussian distribution. Thus, the colloids on the lower plane must be closer to the surface than those on the upper plane.

5. Conclusion

Au colloids have been fabricated by ion implantation into the (001) surface of mica. The surface plasmon polariton was observed at 2.28 eV. A band at 967 cm^{-1} was attributed to a Si-O dangling bond defect. The size of the colloid on different (001) planes is correlated with the Gaussian implantation profile. The use of mica as a substrate for metal colloid fabrication, and for the synthesis of quantum dots by sequential ion implantation[8] when combined with AFM may open new avenues for manipulating nanostructures on an atomically flat, insulating surface.

D. O .H. acknowledges the support from NASA grant No. NAG8-1066 and support from the NASA center of Photonic Materials and Devices The work at ORNL is sponsored by DOE under contract DE-AC05-84OR2 1400 with Lockheed Martin Energy Systems, Inc.

References

1. R. J. Warmack and S. L. Humphrey, Phys. Rev. 334,2246 (1986).
2. K. Fukumi, A. Chayahara, K. Kodono, T. Sakaguchi, Y. Horino, M. Miya, K. Fujii, J. Hayakawa, and M. Satou, J. Appl. Phys. 75, 3075 (1994).
3. W. P. Halperin, Rev. Mod. Phys. 58, 533 (1985).
4. D. O. Henderson, M. A. George, Y. S. Tung, A. Burger, S. H. Morgan, W. E. Collins, C. W. White, R. A. Zuhr, and R. H. Magruder, J. Vat. Sci. Technol. A13, 1254 (1995).
5. E. B. Wilson, Jr., J. C. Decius, and P. C. Cross, The Theory of Infrared and Raman Vibrational Spectra, (Dover Publications, New York 1980) p.65.
6. R. H. Magruder, D. O. Henderson, and R. A. Zuhr, J. Non-Cryst. Solids 152, 258 (1993).
7. D. O. Henderson, R. Mu, Y. S. Tung, M. A. George, A. Burger, S. H. Morgan, C. W. White, R. A. Zuhr, and R. H. Magruder, J. Vat. Sci. Technol. B13, 1198 (1995).
8. C. W. White, J. D. Budai, J. G. Zhu, S. P. Withrow, R. A. Zuhr, Y. Chen, D. M. Hembree, R. H. Magruder, and D. O. Henderson, accepted for publication in the Fall 1994 Materials Research Society Meeting Symposium F: Microcrystalline and and Nanocrystalline Semiconductors, ed. by L. Brus, R. W, Collins, M. Hirose, and G. Koch.

Page intentionally left blank

NASA ... / 63

URC97125

Fuzzy Behavior-based Navigation for Planetary Microrovers

Edward Tunstel*, Harrison Danny, Tanya Lippincott and Mo Jamshidi
NASA Center for Autonomous Control Engineering
Department of Electrical and Computer Engineering
University of New Mexico
Albuquerque, NM 87131

Abstract

Adaptive behavioral capabilities are necessary for robust rover navigation in unstructured and partially-mapped environments. A control approach is described which exploits the approximate reasoning capability of fuzzy logic to produce adaptive motion behavior. In particular, a behavior-baaed architecture for hierarchical fuzzy control of microrovers is presented. Its structure is described, as well as mechanisms of control decision-making which give rise to adaptive behavior. Control decisions for local navigation result from a consensus of recommendations offered only by behaviors that are applicable to current situations. Simulation predicts the navigation performance on a microrover in simplified Mars-analog terrain.

1 Introduction

During the years between 1996 and 2005, NASA will embark on several missions to explore planet Mars. As a part of these exploration initiatives NASA plans to make use of microrovers, which are small mobile robots with mobility characteristics that are sufficient for traversing rough and natural terrain. The first microrover, named Sojourner [1], was launched aboard the Mars Pathfinder spacecraft in December of 1996 and is scheduled to arrive on Mars in July of 1997. This planetary rover is part of the payload of the spacecraft's lander which also carries a stereo imaging system and various science instruments. Sojourner will demonstrate the viability of exploring planetary surfaces using mobile robot technology; its mission will be limited to minimal surface exploration. The focus of ongoing research to develop enabling technology for subsequent microrover deployments is increased mobility and increased autonomy [2, 3]. In this paper, we focus on the latter.

Robustness and adaptability are essential for increasing microrover navigation capabilities beyond those of Sojourner. Realization of robust behavior requires that uncertainty be accommodated by the rover control system. Fuzzy logic is particularly well-suited for implementing such controllers due to its capabilities of inference and approximate reasoning under uncertainty. In order to achieve autonomy, microrovers must be capable of achieving multiple goals whose priorities may change with time. Thus, controllers should be designed to realize a number of task-achieving behaviors that can be integrated to achieve different control objectives. State-of-the-art microrover navigation employs simple behavior control strategies that are based on finite state machines [2, 4]. A different approach which exploits the approximate reasoning facility of fuzzy logic is presented here [5]. It is a hierarchical behavior-based control architecture which enables distribution of intelligence amongst special-purpose fuzzy-behaviors. This structure is motivated by the hierarchical nature of behavior as hypothesized in ethnological models.[1] A fuzzy coordination scheme is

*Jet Propulsion Laboratory, Pasadena, CA.

[1] Models which describe animal behavior patterns.

also described that employs weighted decision-making based on contextual behavior activation. Performance is demonstrated by simulated microrover navigation example in simplified Mars-analog terrain. Interesting aspects of the decision-making process which give rise to adaptive behavior are highlighted.

2 Hierarchical Fuzzy-Behavior Control

The behavior control paradigm has grown out of an amalgamation of ideas from ethology, control theory and artificial intelligence [6, 7]. Motion control is decomposed into a set of special-purpose behaviors that achieve distinct tasks when subject to particular stimuli. Clever coordination of individual behaviors results in emergence of more intelligent behavior suitable for dealing with complex situations. Most behavior controllers have been based on crisp (non-fuzzy) data processing and binary logic-based reasoning [4, 7]. The incorporation of fuzzy logic into the framework of behavior control has been proposed to enhance multiple behavior coordination and conflict resolution [8]. Fuzzy behavior control has also been proposed for autonomous planetary rover navigation in Lunar [9] and Mars [5, 10] missions. Such controllers provide robustness to perturbations, design simplicity, and efficiency in dealing with continuous variables.

In contrast to their crisp counterparts, fuzzy-behaviors are synthesized as fuzzy rule-bases, i.e. collections of a finite set of fuzzy if-then rules. Each behavior is encoded with a distinct control policy governed by fuzzy inference. If X and Y are input and output universes of discourse of a behavior with a rule-base of size n, the usual fuzzy if-then rule takes the following form

$$IF\ \mathbf{x}\ is\ \tilde{A}_i\ THEN\ y\ is\ \tilde{B}_i \tag{1}$$

where x and y represent input and output fuzzy linguistic variables, respectively, and $\tilde{A}_i$ and $\tilde{B}_i (i = 1 \ldots n)$ are fuzzy subsets representing linguistic values of x and y. Typically, x refers to sensory data and y to actuator control signals. The antecedent consisting of the proposition "x is $\tilde{A}_i$" could be replaced by a conjunction of similar propositions; the same holds for the consequent "y is $\tilde{B}_i$".

2.1 Microrover Behavior Hierarchy

In the proposed architecture, a collection of *primitive behaviors* resides at the lowest level which we refer to as the primitive level. These are simple, self-contained behaviors that serve a single purpose by operating in a reactive or reflexive fashion. They perform nonlinear mappings from different subsets of the rover's sensor suite to (typically, but not necessarily) common actuators. Each exists in a state of solipsism, and alone, would be insufficient for autonomous navigation tasks. Primitive behaviors are building blocks for more intelligent composite *behaviors*. They can be combined synergistically to produce behavior(s) suitable for accomplishing goal-directed missions,

Autonomous microrovers must be capable of point-to-point navigation in the presence of varying obstacle (rocks, boulders, dense vegetation, etc.) distributions, surface characteristics, and hazards. Often the task is facilitated by knowledge of a series of waypoints, furnished by humans, which lead to designated goals. In some cases, such as exploration of the surface of Mars [1, 11], this supervised autonomous control must be achieved without the luxury of continuous remote communication between the mission base station and the microrover.[2] Considering these and other constraints associated with planetary rover navigation, suitable behavior hierarchies similar to the hypothetical one shown in Figure 1 could be constructed. In this figure the behavioral functions of goal-seek, route-follow, and localize are decomposed into a suite of primitive behaviors. In Mars exploration mission scenarios [1, 3], microrover position, and all other coordinates of interest, are typically referenced relative to a coordinate frame located at the lander. Thus, any subsequent mention of coordinates or locations refers to the lander coordinate frame of reference. The composite behavior, goal-seek, is responsible for collision-free navigation to a goal location. Route-follow

[2]Time delays between Earth and Mars can be anywhere between 6 and 41 minutes,

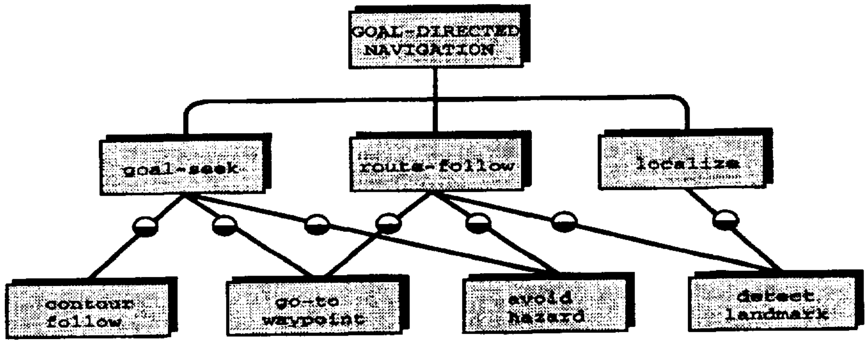


Figure 1: Hypothetical behavior hierarchy for microrover navigation.

is responsible for navigation via a set of waypoints that lead to a goal. Self-localization via dead-reckoning and, perhaps, reference to distinguishable landmarks is the responsibility of localize.

The behavior hierarchy shown implies that goal-directed navigation can be decomposed as a behavioral function of these composite behaviors. They can be further decomposed into the primitive behaviors shown, with dependencies indicated by the adjoining lines. Examples of terrain features which could be considered hazards for microrover navigation include rocks, pits, and excessive slopes. In this paper, we will be primarily concerned wit h rocks. As its name implies, the purpose of the avoid-hazard behavior is to avoid collision with rocks. Later we specify a minimum rock diameter for rocks which are considered hazards. The go-t `o-waypoint` behavior will direct the microrover to traverse a straight line trajectory to a specified waypoint or goal. When close to obstacle (rock) boundaries, cent our-follow maintains the microrover's lateral distance from the obstacle while circumnavigating it. Finally, detect -landmark guides the microrover in search of distinct features which represent landmarks that facilitate self-localization. Interconnecting circles between composite behaviors and the primitive level represent weights and activation thresholds of associated primitive behaviors. Fluctuations in these weights are at the root of the intelligent coordination of primitive behaviors. The hierarchy facilitates decomposition of complex problems as well as run-time efficiency by avoiding the need to evaluate rules from behaviors that do not apply.

Note that decomposition of behavior for a given planetary rover is not unique. Consequently, suitable behavior repertoires and associated hierarchical arrangements are arrived at following a subjective analysis of the system and the task environment. For an actual mission, the design of behaviors at the primitive level would be tailored to the navigation task and an environment with characteristics of natural terrain. The total number, and individual purpose, of fuzzy-behaviors in a given behavior hierarchy is indicative of the problem complexity and can be conveniently modified as required.

3 Coordinating Fuzzy-behavior Interactions

Complex interactions in the form of behavioral cooperation or competition occur when more than one primitive behavior is active. These forms of behavior are not perfectly distinct; they are extremes along a continuum [12]. Coordination is achieved by weighted decision-making and behavior modulation embodied in a concept called the degree *of applicability* (DOA). The DOA is a measure of the instantaneous level of activation of a behavior and can be thought of in ethnological terms as a motivational tendency of the behavior. Fuzzy rules of composite behaviors are formulated such that the DOA, $\alpha_j \in [0, 1]$, of primitive behavior j is specified in the consequent of *applicability rules* of the form

$$IF\ \mathbf{x}\ is\ \tilde{A}_i\ \ THEN\ \ \alpha_j\ is\ \tilde{D}_i \qquad (2)$$

where $\tilde{A}_i$ is defined as in (1). $\tilde{D}_i$ is a fuzzy set specifying the linguistic value (e.g. "high") of α_j for the situation prevailing during the current control cycle. This feature allows certain microrover behaviors to

influence the overall behavior to a greater or lesser degree depending on the current situation. It serves as a form of motivational adaptation since it causes the control policy to dynamically change in response to goals, sensory input, and internal state. Thus, composite behaviors are meta-rule-bases that provide a form of the ethnological concepts of inhibition and dominance. Behaviors with maximal applicability ($\alpha_{max} \leq 1$) can be said to dominate, while behaviors with partial applicability ($O < \alpha < \alpha_{max}$) can be said to be inhibited. These mechanisms allow exhibition of behavioral responses throughout the continuum. This is in contrast to crisp behavior selection which typically employs fixed priorities that allow only one activity to influence the rover's behavior during a given control cycle [4, 7]. The coordination scheme includes behavior selection as a special case when the DOA of a primitive behavior is nonzero and above its activation threshold, while others are zero or below threshold. When this occurs, the total number of rules to be consulted on a given control cycle is reduced. In fact, the number of rules consulted during each control cycle varies dynamically as governed by the DOAs and thresholds of the behaviors involved.

Fuzzy rules of each applicable primitive behavior are processed yielding respective output fuzzy sets, Each fuzzy behavior output is weighted (multiplied) by its corresponding DOA, thus effecting its activation to the level prescribed by the composite behavior. The resulting fuzzy sets are then aggregated and defuzzified to yield a crisp output that is representative of the intended coordination. Since control recommendations from each applicable behavior are considered in the final decision, the resultant control action can be thought of as a consensus of recommendations offered by multiple experts.

4 Microrover Navigation Example

In order to demonstrate fundamental operational aspects of the controller we consider only the composite behavior — route-follow. As illustrated in Figure 1, its effect arises from synergistic interaction between several primitive behaviors. In the following example, avoid-hazard and go-t `o-waypoint` are used. Recall that these behaviors are only capable of exhibiting their particular primitive roles. When more behaviors are involved, the approach proceeds in a straightforward manner by appending additional DO As and any necessary antecedents to applicability rules accordingly. The controller's performance is predicted by simulated microrover navigation in simplified Mars-analog terrain. That is, navigation through a region with a realistic rock distribution is considered, but the terrain is assumed to be two-dimensional. This is an over-simplification of actual microrover mission scenarios in which complex motions in the third dimension occur quite frequently. However, the two-dimensional simplification of Mars-analog terrain still allows us to test the proposed navigation approach in environments densely cluttered by irregularly-shaped obstacles (rocks). Until now, it has only been verified for navigation tasks in indoor office-like environments [5].

The simulated microrover is loosely modeled after *Sojourner*. As shown in Figure 2 its chassis is six-wheeled, with neither axles nor a suspension. It uses a passive rocker bogey mechanism designed to enable climbing over vertical obstacles of 1.5 wheel diameters in height. The $13cm$ diameter wheels are driven by six drive motors; one steering motor is used to independently steer each of the four corner wheels [11]. The steering capability allows for rotating in place. The microrover measures $65cm$ in length, $48cm$ in width and $30cm$ high; its mass is 11 kilograms. Primary navigation sensing consists of light-stripe triangulation (to determine distances to abstacles), turn rate sensing and dead-reckoning (odometry) using wheel encoders. We have simulated the obstacle distance sensing covering an area approximately 1 meter in front of the vehicle, and we have assumed ideal dead-reckoning. Turn rate information was not used. The simulated Martian surface is based on a model of rock size and frequency distributions derived from Viking mission data [13]. The model is know as Moore's model and we have used it here to generate a rock distribution over a $10m$x$10m$ region which replicates the Mars nominal terrain type [13]. The initial state, $(x\, y\, \theta)^T$, of the microrover is $(8.25\ 5.25\ -\frac{\pi}{2})^T$. Its task is to navigate to a goal at (1.75, 6.0) via the following waypoints

$$(7.0, 5.0) \rightarrow (4.0, 6.25) + (2.0, 7.0).$$

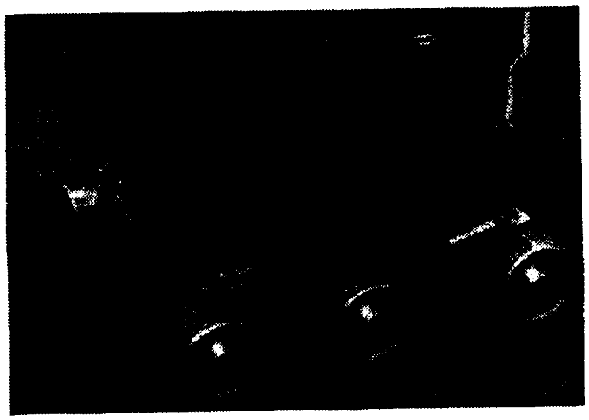


Figure 2: Sojourner: the Mars Pathfinder Microrover.

The resulting route is shown in the left half of Figure 3 where the oddly-shaped icons represent rocks with diameters of $20cm, 40cm$ and $60cm$. Rocks with diameters less than $15cm$ are not considered to be obstacles. The microrover controlled by the fuzzy-behavior hierarchy successfully reaches the goal location via the specified waypoints. In the right half of the figure, the behavioral interaction during the run is shown as a time history of the DOAs of each primitive behavior. The interaction dynamics shows evidence of competition (overlapping oscillations) and cooperation with varying levels of dominance throughout the task. Initially, avoid-hazard has the dominant influence over the microrover. It competes with go-t o-waypoint which dominates as each waypoint is approached. The applicabilities vary continuously reflecting levels of activation recommended by the behavior control system. The individual primitive behaviors are dynamically modulated to produce an overall behavior that accomplishes the navigation objective.

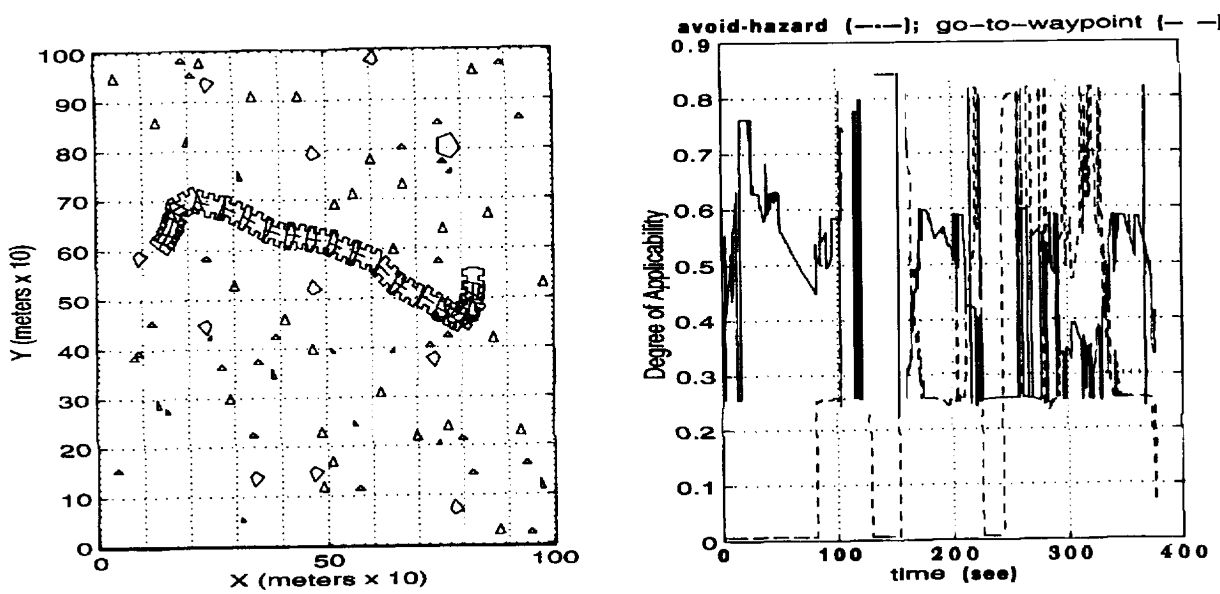


Figure 3: Microrover path and behavior modulation during route-following in Mars nominal terrain.

Acknowledgement:

This work was supported in part by NASA under contract # NCCW-0087.

5 Conclusion

The hierarchy of fuzzy-behaviors provides an efficient approach to controlling mobile vehicles. Its practical utility lies in the decomposition of overall behavior into sub-behaviors that are activated only when applicable. The modularity and flexibility of the approach, coupled with its mechanisms for weighted decision-making, makes it a suitable framework for modeling and controlling situated adaptation in autonomous microrovers. Here, simulation has been used to predict the performance of the approach when applied to microrover navigation in simplified Mars-analog terrain. Successful navigation runs dictate that the approach has potential for applications involving local navigation in densely cluttered, unstructured environments. Future extensions of this work will address three-dimensional simulation in more realistic terrain, and actual experiments pending procurement of a suitable microrover prototype.

References

[1] J. Matijevic and D. Shirley. The mission and operation of the mars pathfinder microrover. In *IFAC 13th* Triennial *World Congress, San* Francisco, CA, 1996.

[2] R. Volpe, J. Balaram, T. Ohm, and R. Ivlev. The rocky 7 mars rover prototype. In *Planetary Rover Technology & Systems Workshop, IEEE International* Conference on *Robotics & Automation,* Minneapolis, MN, April 1996.

[3] S. Hayati. Microrover research for exploration of mars. In *AIAA Forum on Advanced Developments* in Space *Robotics,* Univ. of Wisconsin, Madison, WI, Aug. 1996.

[4] Erann Gat, Rajiv Desai, Robert Ivlev, John Loch, and David Miller. Behavior control for robotic exploration of planetary surfaces. *IEEE Transactions* on *Robotics and* Automation, 10(4):490-503, Aug. 1994.

[5] Edward W. Tunstel Jr. *Adaptive Hierarchy of Distributed Fuzzy Control: Application to Behavior* Control *of Rovers.* PhD dissertation, University of New Mexico, Department of Electrical& Computer Engineering, December 1996.

[6] David J. McFarland. *Feedback Mechanisms in Animal Behavior.* Academic Press, New York, 1971,

[7] Rodney A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation,* RA–2(1):14–23, *1986.*

[8] Edward Tunstel. Coordination of distributed fuzzy behaviors in mobile robot control. In *IEEE International Conference on Systems, Man* and *Cybernetics,* pages 40094014, Vancouver, BC Canada, October 1995.

[9] A. Martin-Alvarez and P. Putz. Interactive autonomy for navigation and piloting of planetary rovers. In *IFAC 13th Triennial World* Congress, San Francisco, CA, 1996.

[10] R. Lea. Fuzzy logic approach to mars rover guidance. In *International* Conference on Fuzzy *Logic and Neural Networks,* Iizuka, Japan, July 1990.

[11] L. Matthies et al. Mars microrover navigation: Performance evaluation and enhancement. *Autonomous Robots, Special* Issue on *Autonomous Vehicles for* Planetary *Exploration,* 2(4):291–311, *1995.*

[12] J.E.R. Staddon. *Adaptive Behavior and Learning.* Cambridge University Press, New York, 1983.

[13] H. Moore and B. Jakosky. Viking landing sites, remote-sensing observations, and physical properties of martian surface materials. Icarus, 81: 164–184, 1989.

URC97126

On Decision-Making Among Multiple Rule-Bases in Fuzzy Control Systems

Edward Tunstel* and Mo Jamshidi
NASA Center for Autonomous Control Engineering
Department of Electrical and Computer Engineering
University of New Mexico
Albuquerque, NM 87131

Abstract

Intelligent control of complex multi-variable systems can be a challenge for single fuzzy rule-based controllers. This class of problems can often be managed with less difficulty by distributing intelligent decision-making amongst a collection of rule-bases. Such an approach requires that a mechanism be chosen to ensure goal-oriented interaction between the multiple rule-bases. In this paper, a hierarchical rule-based approach is described. Decision-making mechanisms based on generalized concepts from single-rule-based fuzzy cent rol are described. Finally, the effects of different aggregation operators on multi-rule-base decision-making are examined in a navigation control problem for mobile robots.

1 Introduction

Many fuzzy controllers proposed in the literature utilize a monolithic rule-base structure. That is, the precepts that govern desired system response are encapsulated as a single collection of *if-then* rules. In most instances, the rule-base is designed to carry out a single control policy or goal. As structure and task constraints are removed from the problem domain, the need for increased system autonomy mandates the development of more sophisticated controllers. Complex intelligent systems must be capable of achieving multiple goals whose priorities may change with time. When employing fuzzy logic, it becomes difficult to formulate monolithic rule-bases which comply with multiple interacting goals, as this requires formulation of a large and complex set of fuzzy rules. In this situation a potential limitation to the utility of the monolithic fuzzy controller becomes apparent. Since the size of complete monolithic rule-bases increases exponentially with the number of input variables [1], multi-input systems can potentially suffer degradations in speed of response. Alternatively, controllers can be designed to realize a number of distributed special-purpose capabilities that can be integrated to achieve different control objectives. This can be done by organizing fuzzy systems into hierarchical rule structures. It has been demonstrated that such rule structures can be employed to overcome the limitation of monolithic structures by reducing the rate of rule increase to linear or piecewise-linear [1]–[2]. Hierarchical rule structures have also been proposed for controlling systems with interacting goals [3].

This paper describes a fuzzy control architecture for complex systems in which distributed intelligence can be represented as hierarchical or decentralized structures, e.g. autonomous mobile vehicles, multi-agent systems, electric power systems, and other large-scale systems. Decision-making mechanisms baaed on generalized concepts from monolithic fuzzy control are described. The effects of different aggregation operators on multi-rule-base decision-making are examined in an example application to a motion control

*Jet Propulsion Laboratory, Pasadena, CA.

problem for mobile robots. These include the following t-conorms: bounded sum, arithmetic maximum, probabilistic sum, and the Sugeno S_λ family [4].

2 Hierarchical Distributed Fuzzy Control

Fuzzy controllers are intelligent control systems that smoothly interpolate between rules, i.e. rules fire to continuous degrees and the multiple resultant actions are combined into an interpolated result. The underlying theory is based on fuzzy sets [5] which are represented by a mathematical formulation known as the membership function. This function gives a degree or grade of membership within a fuzzy set. Over a given universe of discourse X, the membership function of a fuzzy set $\tilde{A}$, denoted by PA(z), maps the elements $x \in X$ into a numerical value in the unit interval, i.e.

$$\mu_{\tilde{A}}(x) : X \rightarrow [0,1]. \tag{1}$$

Within this framework, a membership value of zero corresponds to an element which is definitely not a member of the fuzzy set, while a value of one corresponds to the case where an element is definitely a member of the set. Partial membership is indicated by values between O and 1, continuous. Implementation of a fuzzy controller requires assigning membership functions for both inputs and outputs, thus the membership values are actually measures of degree of causality in an input-output mapping. Inputs to a fuzzy controller are usually measured variables, associated with the state of the controlled plant, that are fuzzified (assigned membership values) before being processed by an inference engine. The heart of the controller inference engine is a *rule-base* of if-then rules whose antecedents and consequences are made up of linguistic variables and associated fuzzy membership functions. Consequences from different rules are numerically aggregated by fuzzy set union operation and are then collapsed (defuzzified) to yield a single real number output that serves as the control signal for the plant.

In our hierarchical approach, each rule-base is encoded with a distinct control policy governed by fuzzy inference. Thus, each rule-base is similar to the conventional fuzzy controller in that it performs an inference mapping from some input space to some output space. If X and Y are input and output universes of discourse of a behavior with a rule-base of size n, the usual fuzzy if-then rule takes the following form

$$IF \text{ x } is \ \tilde{A}_i \ THEN \ y \ is \ \tilde{B}_i \tag{2}$$

where x and y represent input and output fuzzy linguistic variables, respectively, and.& and $\tilde{B}_i$ ($i = 1..\,.n$) are fuzzy subsets representing linguistic values of x and y. Typically, x refers to sensory data or goal information and y to control outputs (inputs to the controlled system). In general, the rule antecedent consisting of the proposition "x is $\tilde{A}_i$" could be replaced by a compound fuzzy proposition consisting of a conjunction (and/or disjunction) of similar propositions. Similarly, the rule consequent "u is $\tilde{B}_i$" could include additional FLC output propositions.

Overall system behavior is decomposed into a bottom-up hierarchy of increased complexity in which activity at a given level is dependent upon activities at the level(s) below. A collection of primitives typically resides at the lowest level which we refer to as the primitive level. These are simple, self-contained sets of rules that serve a single purpose. They perform nonlinear mappings from different subsets of the sensor suite to subsets of control actions. Alone, each primitive rule-base would be insufficient for achieving complex tasks. Primitive rule-bases are building blocks for more intelligent and higher-level competence. That is, they can be combined synergistically to produce composite capabilities suitable for accomplishing goal-directed operations.

Rule-bases at different levels of the hierarchy are generally interconnected. For example, consider the simple two-level structure illustrated in Figure 1 consisting of primitive rule-bases R_{1a} and R_{1b}, and composite rule-base R_{2a}. The interconnection of R_{2a} with the primitives implies that it can be decomposed as a function of the primitives such that the interaction of R_{1a} and R_{1b} produce the desired task-oriented function of R_{2a}.

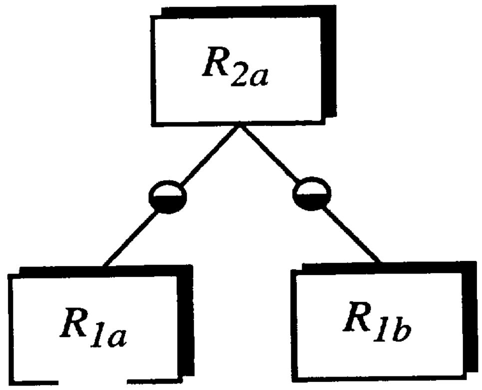


Figure 1: Portion of a rule-base hierarchy.

Note that the overall hierarchy can consist of additional fuzzy rule-bases, the number of which are indicative of the problem complexity. The circles on the adjoining lines in the figure represent weights of the associated primitive rule-bases that fluctuate according to their applicability in the current situation. In general, these weights can be threshold activated. For a given system, the rule-bases and the associated hierarchical arrangement are arrived at following a subjective analysis of the problem and the task domain.

3 Multi-Rule-base Decision-Making

Consensus among multiple rule-bases is achieved using a weighted decision-making strategy embodied in a concept called the *degree of applicability* (DOA). The DOA is a measure of the instantaneous level of activation of a set of rules. The fuzzy rules that make up composite rule-bases are formulated such that the DOA, $\alpha_j \in [0, 1]$, of primitive rule-base j is specified in the consequent of *applicability rules* of the form

$$IF\ x\ is\ \tilde{A}_i\ THEN\ \alpha_j\ is\ \tilde{D}_i \tag{3}$$

where $\tilde{A}_i$ is defined as in (1). $\tilde{D}_i$ is a fuzzy set specifying the linguistic value (e.g. "high") of α_j for the situation prevailing during the current control cycle. This feature allows competence of certain rule-bases to influence the overall system response to a greater or lesser degree depending on the current situation. It serves as a form of adaptation since it causes the control policy to dynamically change in response to goals, sensory input, and internal state. Thus, higher-level rule-bases consist of meta-rules that provide forms of inhibition and dominance. Rule-bases with maximal applicability ($\alpha_{max} \leq 1$) can be said to dominate, while those with partial applicability ($0 < \alpha < \alpha_{max}$) can be said to be inhibited. The hierarchy facilitates decomposition of complex problems as well as run-time efficiency by avoiding the need to evaluate rules from rule-bases that do not currently apply. As a result, the number of rules consulted during each control cycle varies dynamically as governed by the DOAs of the rule-bases involved.

Coordination and conflict resolution are achieved within the framework of fuzzy logic theory –via operations on fuzzy sets [8]. Fuzzy rules of each applicable primitive behavior are processed yielding respective output fuzzy sets. These fuzzy sets are equivalent to the result produced by rule-base evaluation in monolithic fuzzy controllers *before* applying the defuzzification operator. Following consultation of applicable rule-bases, each fuzzy output is weighted (multiplied) by its corresponding DOA, thus effecting its activation to the level prescribed by the higher-level rule-base, The resulting fuzzy sets are then aggregated using an appropriate fuzzy set union operator, and defuzzified to yield a crisp output that is representative of the intended coordination. Since control recommendations from each applicable rule-base are considered in the final decision, the resultant control action can be thought of as a consensus of recommendations offered by multiple experts. This coordination procedure is a generalization of the idea of rule weighting in a monolithic rule base to rule-base weighting among multiple rule-bases. In a similar manner, we use fuzzy set theory to

generalize rule conflict resolution in monolithic rule bases, for resolving conflicts among multiple conflicting rule-bases.

3.1 Aggregation of multi-rule-base outputs

One area of flexibility in this approach is in choosing an appropriate operator for consolidating the multiple control recommendations. We focus on the t-conorm, or generalized union operator of fuzzy set theory. Recall that primitive rule-base outputs are fuzzy sets, and an aggregation across rule-bases must be performed to produce an overall control output. The chosen t-conorm may be different than that used to aggregate the individual *rule* outputs in each rule-base. As the selection of the t-conorm used for rule-base aggregation dictates how anything approaching a consensus will be made, available options should be considered.

We consider the following t-conorms: bounded sum, arithmetic maximum, probabilistic sum, and the Sugeno S_λ family. Their definitions follow, where $\mu_{\tilde{R}}$ and $\mu_{\tilde{S}}$ are membership values describing the fuzzy sets, R and $\tilde{S}$, which are undergoing a union operation. $U(\mu_{\tilde{R}}, \mu_{\tilde{S}})$ denotes the t-conorm, or fuzzy union operator.

Bounded sum:

$$U(\mu_{\tilde{R}}, \mu_{\tilde{S}}) = \min(1, \mu_{\tilde{R}} + \mu_{\tilde{S}}) \tag{4}$$

Maximum:

$$U(\mu_{\tilde{R}}, \mu_{\tilde{S}}) = \max(\mu_{\tilde{R}}, \mu_{\tilde{S}}) \tag{5}$$

Probabilistic sum:

$$U(\mu_{\tilde{R}}, \mu_{\tilde{S}}) = \mu_{\tilde{R}} + \mu_{\tilde{S}} - \mu_{\tilde{R}}\mu_{\tilde{S}} \tag{6}$$

Sugeno S_λ family:

$$U(\mu_{\tilde{R}}, \mu_{\tilde{S}}, \lambda) = \min(1, \mu_{\tilde{R}} + \mu_{\tilde{S}} + \lambda\mu_{\tilde{R}}\mu_{\tilde{S}}) \;\; ; \; \lambda \geq -1 \tag{7}$$

Note that the bounded sum is a special case of the Sugeno family of t-conorms, namely So. Also, the probabilistic sum is very similar to S_{-1}. The S_λ family is one of a variety of parameterized families of aggregation operators; others can be found in [6, 7]. The selection of the above set of t-conorms was based on their computational simplicity (i.e. no division or exponent operations required),

4 Example

For illustration, we apply the hierarchy of distributed rule-bases to a mobile robot navigation example. The task is collision-free and goal-directed navigation from some location to a designated goal inside a hypothetical warehouse environment. A hierarchy of rule-bases for this indoor navigation is arranged as in Figure 2 which implies that goal-directed navigation can be decomposed as a function of goal-seek and route-follow. In the context of mobile robot fuzzy control, these rule-bases are referred to as fuzzy behaviors [8]. These behaviors can be further decomposed into the primitives shown. Avoid-collision and `wall-follow` are self-explanatory. The doorway behavior guides a robot through narrow passageways in walls; `go-to-xy` directs motion along a straight line trajectory to a particular location. In order to demonstrate the decision-making aspects of the controller in the simplest manner possible consider only the composite behavior — goal-seek. As illustrated in Figure 2, its effect arises from synergistic interaction between `go-to-xy` and avoid-collision. The simulated mobile robot is modeled after LOBOt, a custom-built base with a 2-wheel differential drive and two stabilizing casters. It is octagonal in shape, 75 cm tall and 60 cm in width. The output of the primitive behaviors are right and left wheel speeds; the inputs of the hierarchy are the goal location and subsets of sensor readings. The robot's sensor suite includes optical encoders on each driven wheel and 16 ultrasonic transducers arranged primarily on the front, sides, and forward-facing obliques. The initial state of robot is at $(x\, y\, \theta)^T = (11.7m\; 12.3m\; \frac{\pi}{2} rad)^T$. The goal is located at $(1.5m, 1m)$.

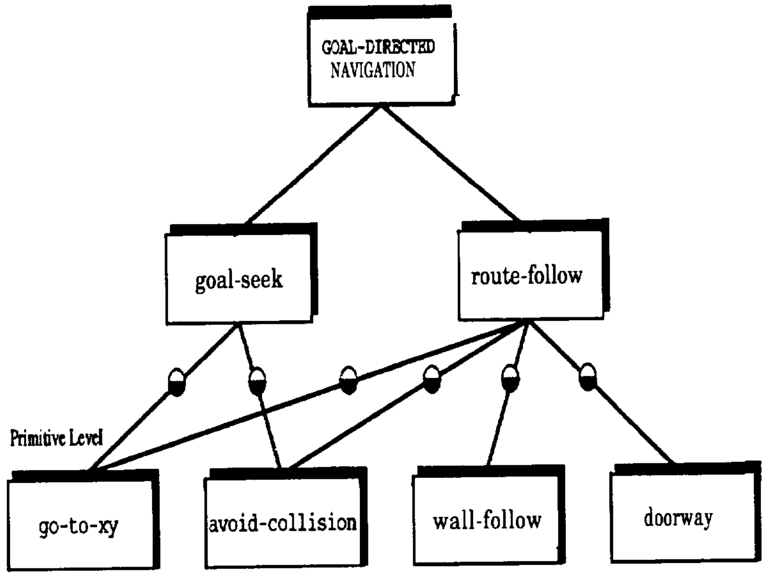


Figure 2: Hierarchical decomposition of mobile robot behavior.

We ran the simulated navigation using each of the t-conorms defined in Section 3.1 to examine the relative impact that each has on motion decisions made during the run. That is, the fuzzy outputs of **go-to-xy** and avoid-collision were aggregated using (2)-(5). The resulting path taken by the robot using the bounded sum t-conorm is shown in Figure 3a; dimensions are meters. The robot simultaneously achieves the goals of reaching the target location and avoiding collisions. The paths resulting from using maximum and probabilistic sum as t-conorms were very similar to the bounded sum case. However, the decisions made as a result of applying the S_λ family for $\lambda \geq 1$ were clearly different as revealed by the alternative path shown in Figure 3b for $\lambda = 1$. In this case, the ensemble of control decisions made over the course of the run have led to a more direct path to the goal. The results were similar for $\lambda > 1$. Thus, possible variations in system behavior can be determined through examination of the effects of t-conorm selection on multi-rule-base decision-making. Of course, the selection of an appropriate t-conorm will be system dependent and baaed on desired system response.

5 Conclusion

Decision-making in distributed intelligent systems is facilitated using a collection of fuzzy rule-based decision systems and controllers. Consensus among multiple rule-bases is achieved using a weighted decision-making strategy based on degrees of applicability y associated with each rule-base. When conditions for activation of a single rule-base (or several) are satisfied, there is no need to consult rule-bases that do not apply. The approach is suitable for fuzzy control of complex, systems that can be represented as hierarchical or decentralized structures. By generalizing decision-making concepts of monolithic fuzzy controllers, it is possible to coordinate multiple distributed rule-bases in a single multi-level control system. When the proposed decision-making mechanisms are employed, it is beneficial to examine different aggregation operators over rule-base outputs to determine the most appropriate operator for the application.

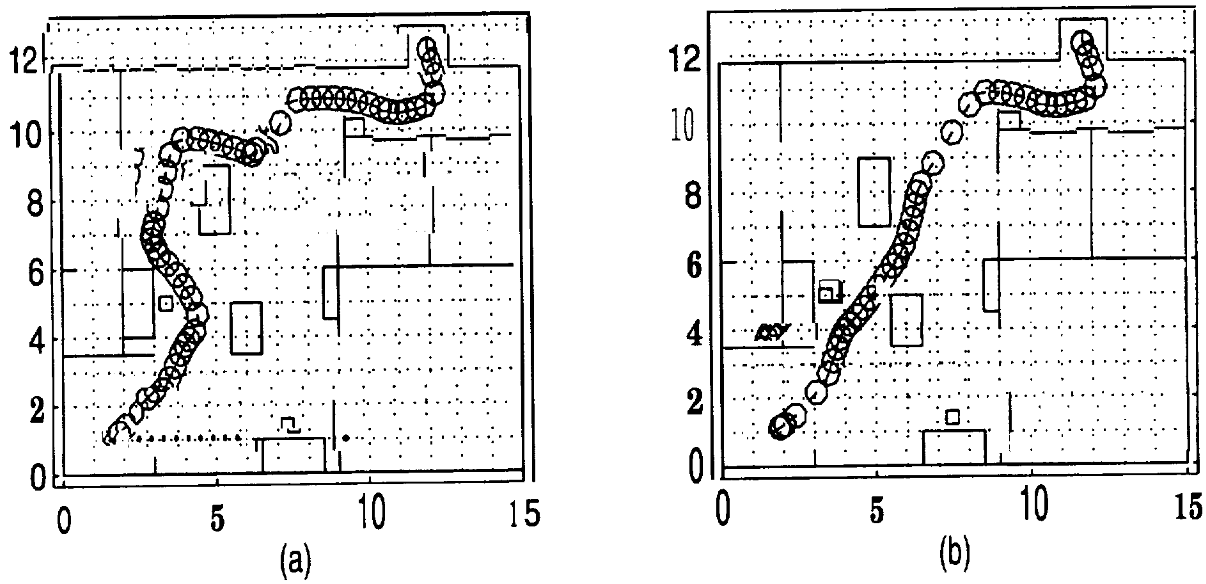


Figure 3: Simulated navigation control using different t-conorms.

References

[1] Raju, G. V. S., J. Zhou and R. A. Kisner, "Hierarchical Fuzzy Control", *International Journal of Control, 54 (5)*, pp. 1201-1216, 1991.

[2] Jamshidi, M. *Large-Scale Systems: Modeling, Control, and* Fuzzy *Logic*, Prentice Hall PTR, Upper Saddle River, NJ, 1997.

[3] Berenji, H.R. et al, "A Hierarchical Approach to Designing Approximate Reasoning-Based Controllers for Dynamic Physical Systems", *6th Conference* on *Uncertainty in AI*, pp. 362-369, 1990.

[4] Driankov, D., H. Hellendoorn and M. Reinfrank *An Introduction to Fuzzy Control*, Springer-Verlag, Berlin, 1993.

[5] Zadeh, L.A. "Fuzzy sets", Information *and Control 12*, pp. 338-353, 1965.

[6] Klir, G.J. and T.A. Folger Fuzzy *Sets, Uncertainty and* Information, Prentice Hall, Englewood Cliffs, NJ, 1988.

[7] Yager, R.R. and D.P. Filev *Essentials of Fuzzy Modeling and Control*, Wiley & Sons, New York, 1994.

[8] Tunstel, E. "Coordination of Distributed Fuzzy Behaviors in Mobile Robot Control", *IEEE* International Conference *on Systems*, Man and *Cybernetics*, Vancouver, BC, pp. 4009-4014, October 1995.

5/27/76

URC97127

Annealing Effects on the Surface Plasmon Of MgO Implanted with Gold

A. Ueda*, R. Mu*, Y.-S. Tung*, D.O. Henderson*(l), C.W. White**, R.A. Zuhr**, Jane G. Zhu**, and P. W. Wang* **

*Chemical Physics Laboratory, Physics Department, Fisk University, Nashville, TN 37208
**Oak Ridge National Laboratory, Solid State Division, Oak Ridge, TN 37830
***Department of Physics, The University of Texas at El Paso, El Paso, TX 79968

ABSTRACT

Gold ion implantation was carried out with the energy of 1.1 MeV into (100) oriented MgO single crystal. Implanted doses are 1, 3, 6, 10 x1 0^{16} ions/cm^2. The gold irradiation results in the formation of gold ion implanted layer with a thickness of 0.2 μm and defect formation. In order to form gold colloids from the as-implanted samples, we annealed the gold implanted MgO samples in three kinds of atmospheres: (1)Ar only, (2)H_2 and Ar, and (3)O_2 and Ar. The annealing over 1200°C enhanced the gold colloid formation which shows surface plasmon resonance band of gold. The surface plasmon bands of samples annealed in three kinds of atmospheres were found to be at 535 nm (Ar only), 524 nm(H_2+Ar), and 560 nm (O_2+Ar). The band positions of surface plasmon can be reversibly changed by an additional annealing.

INTRODUCTION

The reduction of matter to nanometer dimension often leads to significant changes in optical, structural, and thermodynamic properties, The appearance of the surface plasmon (SP) resonance for metal colloid reduced to the nanometer scale is one example of how the optical properties of metals change when they are reduced in size. To form nanometer sized metal particles, we used ion implantation methods [1]. Metal colloid doped glasses, for example, show high optical nonlinearity, and are attractive candidates for utilization in optical devices [2,3]. The enhancement of the third order optical nonlinearity depends on the intensity and frequency of the SP resonance of the metal colloid. According to Mie's scattering theory and effective medium theory [4], we can predict the SP resonance frequency ω_{SP} if we know the dielectric functions of the metal colloid $\varepsilon(\omega)$ and the host material ε_m. As a first-order approximation, the following equation gives us the SP resonance frequency ω_{SP}:

$$\varepsilon(\omega_{SP}) + 2\varepsilon_m = 0. \qquad (1)$$

We have previously reported on the SP resonance of gold nanoparticles in the several substrates: sapphire, CaF_2, and Muscovite mica [5]. MgO is also a material widely used in optics, such as windows and coatings. Although many of the alkaline earth oxides are hydroscopic and thus not widely used in optical systems, MgO is relatively insoluble, hard, and durable. Moreover, MgO is a good transparent material in the range between 300 nm and 7000 run, and the optical absorption

edge has been measured to be about 7.77 eV (160 rim), Transition metals, for example Fe^{3+}, are common multivalent impurities in MgO with the order of 50 ppm. Absorption bands due to Fe^{3+} are found at 210 and 285 nm in MgO [8]. The F and F^{+}absorption bands overlap in MgO, with the F peak occurring at 248 nm and F' peak at 252 nm. In MgO heavily damaged by ion bombardment, a broad absorption feature appears between the F band (-248 nm) and the fundamental absorption edge (-160 rim). Chen et *al.* reported [7] that mobile oxygen interstitial created by electron and neutron irradiation can be annealed out at 400°C, where the interstitial recombine with oxygen vacancies. On the other hand, the oxygen vacancies do not become mobile until the temperature reaches 1000°C [7]. In this paper, we mainly report on how the annealing atmosphere affects to the SP resonance.

EXPERIMENTAL

MgO single crystals (15x15x0.5 mm^3) oriented (100) were obtained from two venders (Harrick Scientific and Commercial Crystal Laboratories). The major impurities are Fe^{3+} of ~100 ppm and Al^{3+} of -40 ppm. The samples are implanted with 1.1 MeV gold $^{197}Au^{+}$ with doses of 1, 3, 6, $10x10^{16}$ Au ions/cm^2. The depth profile of gold concentration was measured by Rutherford Backscattering Spectrometry (RBS), using 2.4 MeV α particles. Thermal annealing was carried out at temperatures between RT and 1200°C in a reducing (10%H_2+90%Ar), an oxidizing (10%O_2+90%Ar), and an inert atmosphere (100%Ar) flow.

The electronic spectra were measured between 185-3200 nm with a UV-Vis-NIR spectrophotometer (Hitachi, U-3 501). In order to study the annealing effects on Au implanted MgO, each transmission spectrum was taken every 15 min annealing, unless otherwise indicated. Indices of refraction of the samples were measured with an ellipsometer (Dudolph Research, 43603 -200E).

RESULTS AND DISCUSSION

The RBS spectra show that the depth profile is, similar to the Gaussian distribution, whose peak of the Au concentration is located about 0.2 µm from the surface (The spectra are not shown here).

The transmission spectra of as-implanted Au/MgO samples and an unimplanted MgO sample (called "virgin") are shown in Fig.1. The virgin sample (a) has a high transmission region (85% transmittance) down to 350 nm, and below 320 nm the transmittance suddenly drops due to Fe^{3+} ion impurities whose absorption bands are located at 285 nm and 210 nm. The spectra (b), (c), (d), and (e) are the transmission spectra for the gold implanted samples with dosages of 1, 3, 6, and 10 $x10^{16}$

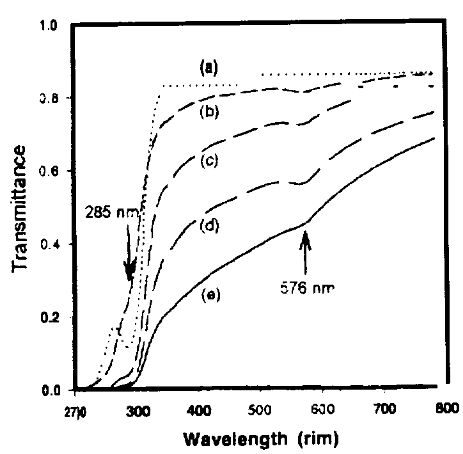


Figure 1 Transmission Spectra of (a) virgin MgO and as-implanted MgO with Au ions of (b)$1x10^{16}$, (c)$3x10^{16}$, (d)$6x10^{16}$, and (c)lx 10" ions/cm^2.

ions/cm^2, respectively. Every spectrum from (b) to (e) has a broad absorption band at 576 nm, which can be attributed to F.-center (aggregates of F-centers) due to the irradiation damage by ion implantation [6]. As the gold dosage increases, the transmission decreases in the range between 300 nm and 1000 nm, mainly due to defects and scattering from small gold particles. The color of the as-implanted samples is dark brownish.

It has been reported that the bands around 250 nm and 355 nm are related to F-centers and F^+-centers caused by neutron-irradiation, electron-irradiation and Mg-vapor-deposition [7]. The bands caused by neutron- and electron-irradiation can be annealed out at 600°C, while these bands generated by Mg-vapor-deposition could not be annealed out until the annealing temperature reaches 1200°C. The reason for this is as follows: in the sample, excessive Mg^{2+} ions are introduced from Mg vapor and then the Mg ions migrate into the MgO bulk to cause oxygen ion vacancies, In this case there exist no interstitial oxygen ions in the system, while neutron- or electron- irradiated samples have interstitial oxygen ions to re-combine to oxygen vacancies, The oxygen interstitial are mobile, while oxygen ion vacancies *are* immobile [7]. Below 300 nm, the Fe^{3+} band is overlapped with the F- and F^+-band, thus it is difficult to distinguish those bands in the as-implanted samples before annealing.

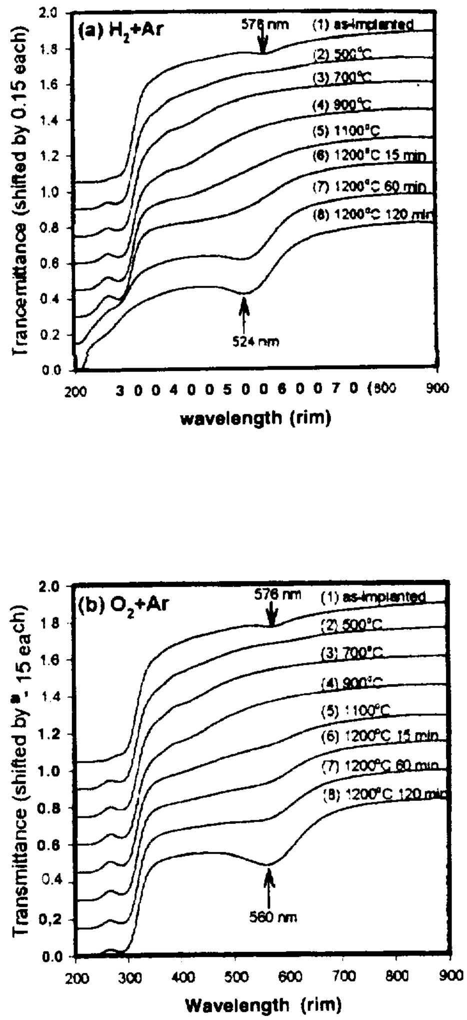


Figure 2 Transmission spectra of the sequentially annealed Au/MgO in (a)H_2+Ar and (b)O_2+Ar.

In Fig.2, two sets of spectra for sequentially annealed Au/MgO sample in both reducing and oxidizing atmospheres are shown, for the sample with the dosage of 6x10^{16} ions/cm^2. The annealing temperature and time for each spectrum are shown in Fig.2. The band at 576 nm disappeared after annealing at 400°C, as reported by Chen *et al.* [7]. Between 200°C and 1000°C, weak broad bands appeared around 340 and 365 nm. We will discuss these bands elsewhere. When the annealing temperature reached 1200°C, a strong broad band appeared at 524 nm for the samples annealed in the reducing atmosphere and at 560 nm for the oxidizing atmosphere, which are attributed to the SP resonance of gold colloids. As the annealing time increased at 1100- 1200°C, the SP bands became stronger, which indicates that at this temperature (1 100- 1200°C) the implanted gold *ions* begin to diffuse and aggregate to form metallic colloids. At the same time of colloid formation, the Fe^{3+} band at 285 nm became weaker in the case of reducing atmosphere. It is also coincident that the temperature at which oxygen vacancies become mobile, as mentioned before.

In order to investigate the correlation between the SP resonance and the reduction of Fe^{3+} band at 285 nm, we measured the transmission spectra of virgin MgO that was sequentially annealed in the similar way to the previously gold implanted samples. In Fig.3, (a) spectrum for virgin MgO is shown before annealing. Spectra (b), (c), and (d) are of the same sample after sequential annealing at 1000°C, 1100°C, and 1200°C, respectively, for 15 min. Spectra (e) to (i) are of the same sample additionally annealed at 1200°C for 15 min for each step. During annealing between 200°C and 900°C with a step of 100°C for 15 minutes each, the spectra did not exhibit any significant changes. After annealing at 1200°C for 90 minutes in total, (i) in Fig.3, the surface of the sample became slightly cloudy. However, before it became cloudy, the Fe^{3-} band at 285 nm became weak, Apparently,the Fe^{3+} ions were reduced to Fe^{2+} ions in the reducing atmosphere at high temperature, which resulted in an increased transmittance of 75°/0 at 285 nm [8]. After annealing in a reducing atmosphere, we switched the atmosphere to O_2+Ar. The transmittance suddenly dropped just after annealing at 1200°C in oxidizing atmosphere for 15 minutes, as shown in Fig.3(b). Since we can assume that Fe^{3+}ions uniformly distributed in the entire sample, the reducing reaction occurs from surface to inside the sample until Fe^{3-} ions are almost completely converted to Fe^{2-} ions in reducing atmosphere. When annealing the samples in oxidizing atmosphere, the conversion from Fe^{2+} to Fe^{3-} has taken place, as shown in Fig.3(b). However, this does not mean that both hydrogen and oxygen atoms diffuse into the sample. In the reducing atmosphere at high temperature, a hydrogen molecule takes an oxygen atom at the surface, leaving an oxygen vacancy with two electrons behind and giving one electron to an Fe^{3+} ion to form an Fe^{2+} ion, This reaction occurs only at the surface and will not continue inside the sample at lower temperature because oxygen vacancies are not mobile at low temperature (<1000°C). This is consistent with our observation that the spectra Fe^{3+} absorption decreases only at high temperature (≥1200°C). In the oxidizing atmosphere, oxygen vacancies are filled with oxygen and take electrons from Fe^{2+} to make them Fe^{3+} thereby allowing oxygen to diffuse into the sample when the temperature is high enough.

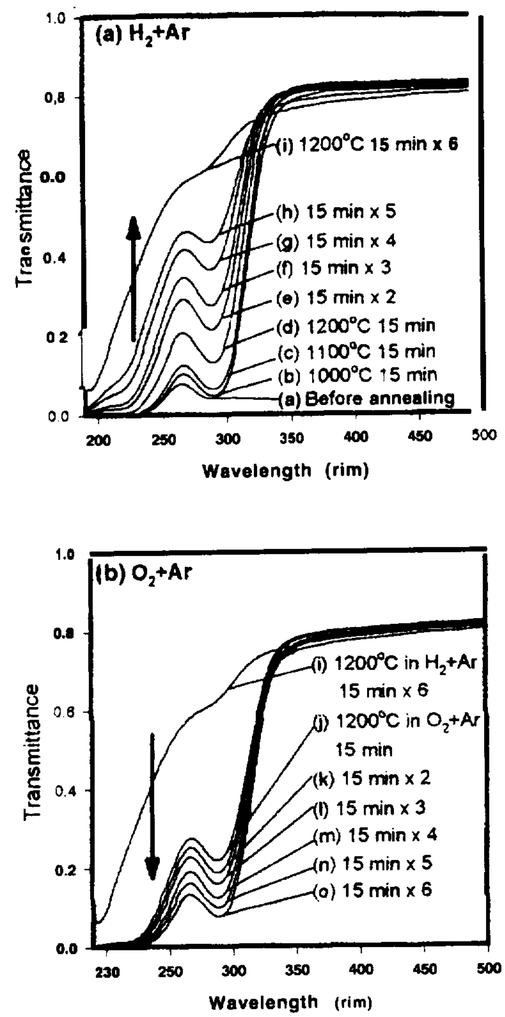


Figure 3 Transmission spectra of virgin MgO annealed in (a)H_2+Ar and (b)O_2+Ar in the range for Fe^{3+} absorption band.

To confirm the annealing atmosphere dependence of the SF band position, we sequentially annealed an implanted sample (Au/MgO with dose of $10x10^{16}$ inos/cm^2) at 1200°C in (1)Ar only, (2)O_2+Ar, (3)H_2+Ar, and (4)O_2+Ar again for 2 hours each. Fig.4 shows the vicinity (a) of the SP band and wider range of spectra in the inset (b). (1)After annealing in Ar only, SP appeared at 535 nm. (2)After annealing in O_2+Ar, SP shifted to 560 nm. (3)Annealing in H_2+Ar made SF shift to 524 nm. (4) Again, after annealing in O_2+Ar, SP moved back to 560 nm. From these results, the SP position is reversibly depending on the annealing atmosphere. This reversibility suggests that theSP band position is influenced by the local dielectric constant of host near gold colloids. Once colloids with a certain size were formed, it is difficult to change their size or shape. Using these values of ω_{SP} from experimental data, Eq. (1), and the optical constants of gold from literature [9], we calculated the local ε_m to be 3.01 (Ar only), 2.45 (H_2+Ar), and 3.90 (O_2+Ar), while the standard value of ε_m for MgO is 3,02.

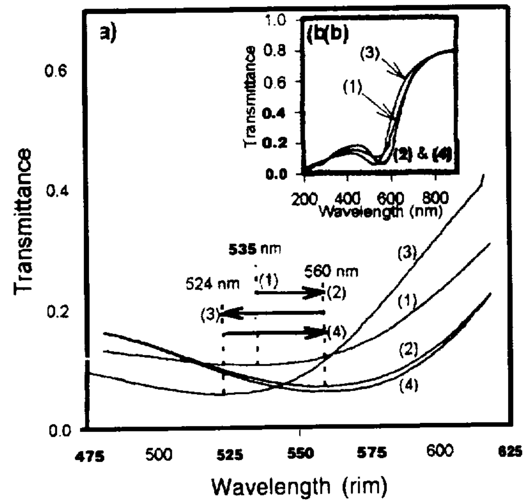


Figure 4 Shifts of surface **plasmon** bands due to **different** annealing atmospheres: (1)Ar only, (2)O_2+Ar, (3)H_2+Ar, and (4)O_2+Ar at 1200°C for 2h each.
Inset (b) shows wide range of spectra.

CONCLUSIONS

We have found that the annealing atmosphere affects the position of SF band of Au implanted in MgO. According to the temperature dependence of SF resonance development, oxygen vacancy mobility and reduction of Fe^{3+} ions, the gold colloid formation must be well related to oxygen vacancy movement at high temperature. The shift and the reversibility may be useful for application to the development of nonlinear optical devices. Further study is needed to understand the mechanismof the annealing effects on surface plasmon bands.

ACKNOWLEDGEMENTS

We gratefully acknowledge research that was supported by NASA grant No. NAG8-1 066. The research at ORNL was sponsored by hte Division of Materials Sciences, U. S. Department of Energy, under Contract No. DE-AC05-9OR22464 with Lockheed Martin Energy Systems and the NASA Center for Photonic Materials and Devices.

REFERENCE

[1] C. W. White *et al.*, Mat, Science Reports, 4, 43(1989).

[2] K. Fukumi, *et al.*, J. Appl. Phys. 75, 3075(1994),

[3] R. F. Hugland *et al.*, Nucl. Instr. and Method, **B65**, 405(1992).

[4] J. A.A.J. Perenboom, P, Wyder, and F, Meier, Phys. Rep. 78, 173(1981).

[5] D. O. Henderson *et al.*, J. Vat. Sci. Technol. **B 13**, 1198(1995); D. O. Henderson *et al., J.* Non-Cryst. Solids (to be published); D. O. Henderson *et al.*, J. Vat. Sci. Technol. **A 14**, 1199(1996).

[6] *"Handbook of Laser Science and Technology,"* (CRC Press, Boca Raton, Florida 1986)

[7] Y. Chen, R.T. Williams, and W.A. Sibley, Phys. Rev. 182, 960 (1969).

[8] A. Briggs, J. Mater. Sci. 10, 737(1975).

[9] *"Handbook of Optical Constants of Solids* I & II," Ed. by E.D. Palik, (Academic Press),

URC97128

ACCELERATION OF A FUZZY CONTROLLER USING THE CHINESE REMAINDER THEOREM

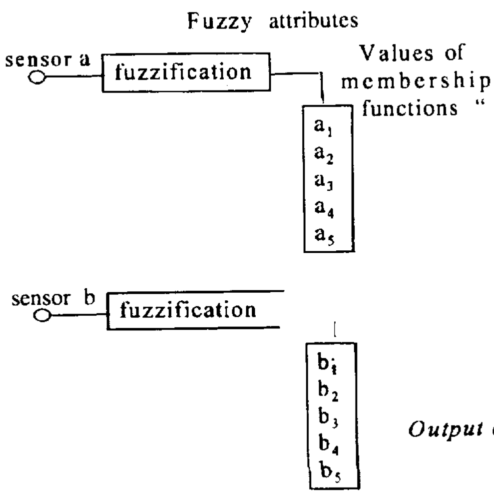
F. Vainstein, telephone: (910) 334-7760 x222, email: feodor@ncat.edu
C. V. Curtis, telephone: (910) 334-7761 x259, email: cvc@ncat.edu
The NASA Autonomous Control Engineering Center
Department of Electrical Engineering, North Carolina A& T State University

Abstract - The paper presents a new approach to increasing the speed of a fuzzy controller. We propose to use a computation algorithm based on the Chinese Remainder Theorem. We will also show how Genetic Algorithms can be used for hardware minimization.

keywords: fuzzy controller, fast adders, fast multipliers, chinese remainder theorem, modular arithmetic, genetic algorithms, hardware design, optimization methods, control theory.

I. INTRODUCTION

The work of a fuzzy controller is demonstrated by the example shown in Figure 1 [1].



POSITION

VELOCITY	NM	NS	ZE	PS	PM
NM	1	4	3	2	1
NS	5	2	1	3	2
ZE	1	3	2	1	4
PS	5	5	3	2	1
PM	1	2	5	1	2

RULE BASE

Output of a fuzzy controller $y = \dfrac{\sum_{i,j=1}^{5} a_i b_j y_{ij}}{\sum_{i,j=1}^{5} a_i b_j}$

Figure 1. **Fuzzy Controller.**

The input of the controller is the position and velocity of the controlled object. The input information is obtained by sensors a and b. The crisp values, a and b, are fuzzified and the values of the membership functions are denoted by $a_1, \ldots a_5$, and $b_1, \ldots b_5$. The output of the fuzzy controller is computed using

the following formula: $y = \frac{\sum_{i,j=1}^{5} a_i b_j y_{ij}}{\sum_{i,j=1}^{5} a_i b_j}$, where the numbers y_{ij} are obtained from the Rule Base table.

We want to point out that the computation of y has to use the operations of addition and multiplication several times, while the operation of division is used only once. The quantity of calculations grows exponentially with the number of sensors in the controller. Therefore, the problem of speeding up the computational process becomes increasingly important. Traditional methods of speeding up computations, due to their generality, will not provide **us** with the fastest possible calculations [2]. We propose to use modular arithmetic based on the Chinese Remainder Theorem (CRT).

II. CHINESE REMAINDER THEOREM (CRT)

The proposed method of computations is based on the Chinese Remainder Theorem. The interested reader can find more details about this theorem in [3].

Theorem: Let be ideals of A. Assume that $\alpha_i + \alpha_j = A$ for $i \neq j$.

Let $$f: A \to \prod_{i=1}^{n} A/\alpha_i = (A/\alpha_1) \times \cdots \times (A/\alpha_n)$$

be a mapping of A into the product induced by the canonical map of A onto A/α_i for each factor. Then

the kernel of f is $\bigcap_{i=1}^{n} \alpha_i$, and f is subjective, thus giving an isomorphism $A \bigcap A/\alpha_i \xrightarrow{\approx} \prod A/\alpha_i$.

We are going to use the following corollary of this theorem.

Corollary: Let $m_1, \ldots, m_n$, be integers relatively prime in pairs. Then

a) any integer k, $O \leq k \leq M$, where $M = (m_1 \cdot m_2 \cdot \ldots \cdot m_k) - 1$ is uniquely represented by a string of n numbers $k \bmod m_1, \ldots, k \bmod m_n$;

b) this presentation preserves operations of componentwise addition and multiplication,

<u>Example</u> Let $m_1 = 2$; $m_2 = 3$; $m_3 = 5$; $m_4 = 7$;

The integers from O to $M = (2.3.5.7) - 1 =$ ***209 can*** be uniquely represented. Table 1 shows the representation of the integers from O through 15.

	mod 2	mod 3	mod 5	mod 7
0	0	0	0	0
1	1	1	1	1
2	0	2	2	2
3	1	0	3	3
4	0	1	4	4
5	1	2	0	5
6	0	0	1	6
7	1	1	2	0
8	0	2	3	1
9	1	0	4	2
10	0	1	0	3
11	1	2	1	4
12	0	0	2	5
13	1	1	3	6
14	0	2	4	0
15	1	0	0	1

Table 1. CRT Number Representation Table

According to Table 1, 3++ 1033, $4 \mapsto 0144$. Let us perform the addition and multiplication of the numbers 3 and 4 using the table above.

$3+4=7 \mapsto 1120;$

$1+0 \bmod 2=1;\;\; O+1 \bmod 3=1;\;\; 3+4 \bmod 5= \;2; 3+4 \bmod 7=0;$

$3 \cdot 4=12 \mapsto 0025;$

$1 \cdot 0 \bmod 2=0;\;\; 0 \cdot 1 \bmod 3=0;\;\; 3 \cdot 4 \bmod 5=2;\;\; 3 \cdot 4 \bmod 7=5;$

Note that the additions and multiplications are done componentwise. The calculations can be done in parallel, because no carry has to be generated. The idea of using CRT for computations is not new [4]. The disadvantage of this approach is that the operation of division is not easily performed. However, a fuzzy controller has to execute the operation of division only once, so for division traditional methods can be used.

The problem of converting CRT numbers to and from binary (or decimal) can be easily solved. Let us denote the function that performs conversion by $\varphi(k)$: q(k)= $(k \bmod m_1, \ldots, k \bmod m_n)$

It is clear that $\varphi(c_0+c_1 \cdot 10+ \ldots +c_d \cdot 10^d)=\varphi(c_0)+\varphi(10) \cdot \varphi(c_1)+\cdots+\varphi(c_d) \cdot \varphi(10)^d$ and this formula can be used for converting from decimal to CRT numbers. Converse] y, a CRT number $a_1, \ldots, a_k$ can be presented by $\sum_{i=1}^{k} a_i e_i$, where $e_i = 0, \ldots, \underset{i}{1}, \ldots, 0$ (the string consists of all O's, except one 1 in the *ith*

position). The inverse transform, φ^{-1}, can be determined according to the following equation:

$\varphi^{-1}\left(\sum_{i=1}^{k} a_i e_i\right) = \sum_{i=1}^{k} \varphi^{-1}(a_i) \cdot \varphi^{-1}(e_i)$, where the numbers of $\varphi^{-1}(e_i)$ can be precomputed.

We want to acknowledge that the CRT numbers can be obtained by using specialized Analog-to-Digital converters in the initial phase of data processing by a fuzzy controller. The conversion of CRT numbers to binary (or decimal) is performed only once in the final stage of the computation for division.

Let us compare the speed of CRT adders/multipliers with the traditional fast adders/multipliers. The CRT adder/multiplier is, basically, a set of modulo m_i adders/multipliers. Even for the case of relatively small m_i, a significant value of m can be obtained. For example, if $\{m_i\} = \{1\ 1,13,17,19,23,25,27,28,29,31\}$, then $M \approx 1.8 \cdot 10^{13} \approx 2^{44}$ Modulo m_i adders/multipliers can be implemented using two or three gate level circuits. The result of this comparison is shown in Table 2 (for a two gate level implementation).

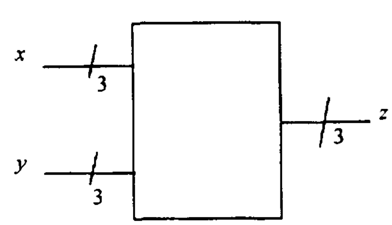
	Time to perform operation	Time to perform operation for n=32
Carry Lookahead Tree Adder	$2\lceil \log_2 n \rceil \tau$	10τ
Wallace Tree Multiplier	$2\lceil \log_2 n \rceil^2 \tau$	50τ
CRT Adder	2τ “	2τ
CRT Multiplier	2τ	2τ

n = number of bits, τ = average delay on u gate

Table 2. Comparison of Chinese Remainder Adders/Multipliers with Traditional Fast Adders/Multipliers

III. HARDWARE IMPLEMENTATION

The problem of hardware minimization can be demonstrated by a modulo 5 adder. The numbers { O, 1, 2,3, 4 } has to be represented in binary; three bits should be allocated for each individual number. A modulo 5 adder is shown in Figure 2.



$(z_1, z_2, z_3,$

Figure 2 Modulo 5 Adder

With the conventional number representation, the numbers O is presented by 000, the number 1 by 001, ...the number 4 by 100 and us have the following truth table for z_1, z_2, and z_3.

x2	y1	y3	z2
0	0	0	0
0	0	1	0
0	0	0	1
0	0	1	1
0	1	0	0
0	x	x	x
0	0	0	0
0	0	1	1
0	0	0	1
0	0	1	0
0	1	0	0
0	x	x	x
1	0	0	1
1	0	1	1
1	0	0	0
1	0	1	0

x2	y1	y3	z2
1	1	0	0
1	x	x	x
1	0	0	1
1	0	1	0
1	0	0	0
1	0	1	0
1	1	0	1
1	x	x	x
0	0	0	0
0	0	1	0
0	0	0	0
0	0	1	1
0	1	0	1
0	x	x	x
x	x	x	x

Table 3 Truth Table of the Modulo 5 Adder with conventional number representation.

Using Espresso [5], a hardware minimization tool, we found that with the conventional number representation we need eighteen AND gates and three OR gates. Comparatively, our experiments show that if we use a non-traditional number representation we can achieve at least a twenty percent reduction in the amount of gates needed.

By definition a number representation is an injective mapping

$f: \{0,1,2,3,4\} \to \{000,001,010,...,111\}$.

We denote by S the set of all such functions. Infectivity means that if $x \neq y$, then $f(x) \neq f(y)$.

Clearly the total number of injective functions in this case is equal to $\binom{5}{8}$ .5 ! = 6,720. Every number representation f uniquely defines three Boolean functions $z_1 f), z_2(f)$, and $z_3(f)$. Let us introduce the function F, which for every Boolean function g $Z_2^n \to Z_2$, where $Z_2 = \{O, 1\}$, gives its number of product terms in the minimum sum of products representation. Let us also define a fitness function ψ: S $\to N$, where N is the set of natural numbers, by the following formula.

$$\psi(f) = F(z_1(f)) + F(z_2(f)) + F(z_3(f))$$

[n other words, $\psi(f)$ is equal to the number of AND gates for the adder implementation if number representation f is used.

Genetic Algorithms can be used for minimization of the function ψ . We propose the following structure of a chromosome. A chromosome has two parts H and T, which can be called the head and tail, respectively. Let us denote 000 by $\overline{0}$, 001 by $\overline{1}$, ... and 111 by $\overline{7}$. The head consist of the set of nonused binary strings: $H = \{\overline{0}, \overline{1}, ... , \overline{7}\} - \{f(0), f(1), \ldots, f(4)\}$. The tail, T, is a permutation corresponding to the particular For example, the chromosome corresponding to the following number

representation $f(0) = 101$, ~(1)= 001, $f(2) = 111$, $f(3) = 010$, $f(4) = 110$ has $H = \{\bar{0}, \bar{1}, \ldots, \bar{7}\} - \{\bar{5}, \bar{1}, \bar{7}, \bar{2}, \bar{6}\} = \{\bar{0}, \bar{3}, \bar{4}\}$. The tail T = (3,1,5,2,4) shows the order in which the binary triples are used. Therefore, the chromosome, for this example, is $\{\bar{0}, \bar{3}, \bar{4}\}(3, 1,5,2, 4)$. The order is not essential for the head, but is very essential for the tail.

The operation of mutation is defined separately for the head and tail. The mutation on the head is defined as follows. One entry from the head is replaced by a number from the set of $\{\bar{0}, \bar{1}, \ldots, \bar{7}\}$, which does not belong to the head. As a result of mutation, for example, the chromosome $\{\bar{0}, \bar{3}, \bar{4}\}(3, 1,5,2, 4)$ can be changed to $\bar{5}, \bar{3}, \bar{4}\}$ (3, 1,5,2, 4). Mutation on the tail can be defined as a simple transformation where we trade two entries. For example, $\{\bar{0}, \bar{3}, \bar{4}\}(3,1,5,2,4)$ will be changed to $\{\bar{0}, \bar{3}, \bar{4}\}(3, 4,5,2,1)$.

The crossover operation is defined the following way. Suppose we have two chromosomes, $\{h_1^1, h_2^1, h_3^1\}(t_1^1, t_{21}^1, t_3^1, t_4^1, t_5^1,) = (H^1, T^1)$ and $\{h_1^2, h_2^2, h_3^2\}(t_1^2, t_{21}^2, t_3^2, t_4^2, t_5^2,) = (H^2, T^2)$. These two parents will always have two offspring (HI, T) and (172, T). The children have different heads, which they inherited from their parents, and the same tail, which is obtained from a probabilistic election process that we shall now describe.

We first find t_1. If $t_1^1 = t_1^2$, then $t_1 = t_1^1$. If $t_1^1 \neq t_1^2$, then t_1 takes the values of t_1^1 or t_1^2 with equal probability. Now, suppose that we have already assigned the value for *t*,,..., t_i. The value of t_{i+1} is defined as follows. If neither t_{i+1}^1 nor t_{i+1}^2 belong to the set $\{2,,t_i\}$, then t_{i+1} is assigned to t_{i+1}^1 ‘ r t_{i+1}^2 with equal probability. If either t_{i-1}^1 or t_{i+1}^2 belongs to the set $\{t_1, \ldots, t_i\}$, say t_1^1, then $t_{i+1} = t_{i+1}^2$ (the other element). If both t_{i+1}^1 and t_{i+1}^2 belong to the s{t/t,,...,t_i}hen t_{i+1} *is* randomly selected with equal probability from the set $H = \{1, \ldots, 5\} - \{t_1, \ldots, t_i\}$.

The results of the computer minimization are not available at the present moment. However, we plan to present the results in future publications.

IV. CONCLUSION

In this paper, we have shown that the speed of a fuzzy controller can be increased, considerably by using the Chinese Remainder Theorem number representation. The problem of hardware minimization has been formalized. This gives us the opportunity to use Genetic Algorithms. Due to the specifics of the optimization problem, the operations of mutation and crossover have been defined nontraditionally.

REFERENCES

[1] A. Homaifar and E. McCormick, "Simultaneous Design of Membership Functions and Rule Sets for Fuzzy Controllers Using Genetic Algorithms," *IEEE Trans. Fuzzy Syst.*, vol. 3,no.2, pp.] 29-139, May 1995.

[2] J. L. Hennessey and D. A. Patterson, *Computer architecture: a quantitative approach. San* Francisco: Morgan Kaufmann, 1995.

[3] S. Lang, *Algebra.* Reading, MA: Addison-Wesley, 1993.

[4] L. L. Dornhoff and F. E. Hohn, *Applied Modern Algebra. New* York: Macmillan, 1978.

[5] R. Rudell, A. Sangiovanni-Vincentelli,"Espresso-MV: Algorithms for Multiple-Valued Logic Minimization," *Proc. Cust. Int. Circ. Conf.*, Portland, May 1985,

URC97129

Utility of BRDF Models for Estimating Opt imal View Angles in Classification of Remotely Sensed Images *

P. F. Valdez[a] and G. W. Donohoe[a]

[a]NASA ACE Center for Autonomous Control Engineering
Electrical and Computer Engineering
University of New Mexico
Albuquerque, New Mexico 87131-1356 USA
pvaldez@eece.unm.edu donohoe@eece.unm.edu

1. Introduction

Statistical classification of remotely sensed images attempts to discriminate between surface cover types on the basis of the spectral response recorded by a sensor. It is well known[1-3] that surfaces reflect incident radiation as a function of wavelength producing a spectral signature specific to the material under investigation. Multispectral and hyperspectral sensors sample the spectral response over tens and even hundreds of wavelength bands to capture the variation of spectral response with wavelength. Classification algorithms then exploit these differences in spectral response to distinguish between materials of interest. Sensors of this type, however, collect detailed spectral information from one direction (usually nadir); consequently, do not consider the directional nature of reflectance potentially detectable at different sensor view angles.

Improvements in sensor technology have resulted in remote sensing platforms capable of detecting reflected energy across wavelengths (spectral signatures) and from multiple view angles (angular signatures) in the fore and aft directions. Sensors of this type include: the moderate resolution imaging spectroradiomet er[4] (MOD IS), the multiangle imaging spectroradiometer[5] (MIS R), and the airborne solid-state array spectroradiometer[6] (AS AS).

A goal of this paper, then, is to explore the utility of BRDF models in the selection of optimal view angles for the classification of remotely sensed images by employing a strategy of searching for the maximum difference between surface BRDFs. After a brief discussion of directional reflect ante in Section 2, attention is directed to the Beard-Maxwell BRDF model and its use in predicting the bidirectional reflectance of a surface. The selection of optimal viewing angles is addressed in Section 3, followed by conclusions and future work in Section 4.

1.1. NEFDS Spectral Database

A collection of spectral datasets was obtained from the National Imagery Resource Library (NIRL) and used in this study. Intimately related to the Materials Exploitation Database (MED), the Spectrum Archival Library (SAL), and the Spectral Catalog,[7] the Nonconventional Exploitation Fact ors Data System (NEFDS) database is composed of spectral reflect ante measurement data in the visible/near-infrared (VNIR), the mid-infrared (MIR) and combined visible/near-mid-infrared (VNMIR) for samples of selected materials.

Spectral samples of a concrete runway, a galvanized steel rooftop, and a couple of painted surfaces in the visible/near-infrared (VNIR) were selected for input to a BRDF model for the pairwise estimation of optimal view angles in the classification of remotely sensed images. Detailed properties of the materials used in this analysis have been assembled and are presented in the Spectral Catalog cited earlier.

*This work was supported in part by NASA ACE under contract #NCCW-0087.

2, Directional Reflectance

The *directional-hemispherical* reflect ante defined as

$$\rho'_{dh} \cdot \frac{d\Phi_r}{d\Phi_i} \tag{1}$$

is the fraction of incident radiant flux density from direction, (θ_i, ϕ_i) that is reflected by the surface into all possible directions of the hemisphere surrounding the sample. By definition, then, ρ'_{dh} does not provide any information on the directional nature of reflectance and is only adequate for surfaces that reflect isotropically.

Many surface materials of interest, however, are anisotropic reflectors so that the spectral response is also dependent on the viewing geometry of the sensor.[8–10] This fact has lead to classification difficulties in which single view imagery was used to discriminate between surface materials.[11] A more general and useful description for the directional nature of reflected radiant flux is contained in the *bidirectional reflectance distribution function* (BRDF) which relates the directional distribution of exitant flux to the incident radiant flux striking the surface.

2.1. Bidirectional Reflectance

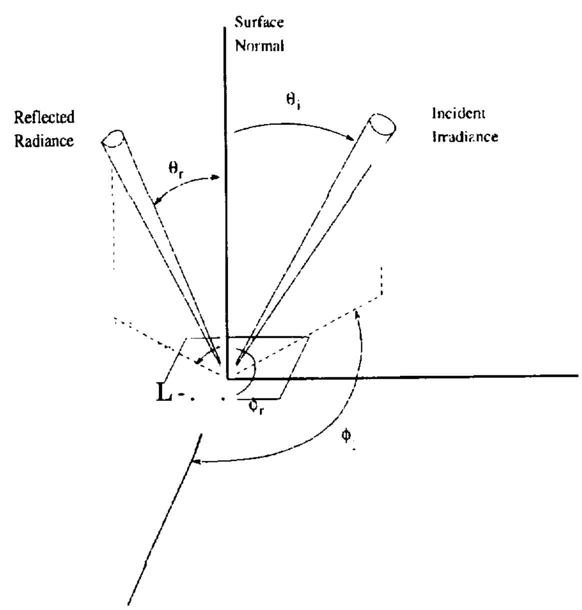


Figure 1. Co-ordinate System Defining Solar Angles and Sensor Viewing Angles.

The BRDF is a function of four angles (Figure 1) and is defined as the ratio of reflected radiance from a surface to the irradiance incident to the surface from an illuminating source. This relationship is seen in Equation 2

$$\rho'_{bd}(\Theta_i, \Theta_r; \lambda) = \frac{dL_r(\Theta_i, \Theta_r; \lambda)}{dE_i(\Theta_i; \lambda)} \tag{2}$$

where the dependence on incident angles, $\Theta_i = \{\theta_i, \phi_i\}$, exit ant angles, $\Theta_r = \{9,, \phi_r\}$, and wavelength, λ, is shown.

2.2. BRDF Models

Reflection of light from a surface is a complex phenomenon which can be very difficult to explain and accurately predict. Even with a firm grasp of the physics involved, the dynamics of a changing environment complicates matters so that a complete understanding is often not possible. As a result, several types of BRDF models have emerged; those based on first principles such as radiative transfer theory,[12–14] geometrical optics ,[15,16] physical optics (wave theory of light) and those that fit analytic equations to the observed reflectance.[17]

Two mechanisms of reflect ante are generally identified: surface and subsurface (volumetric) reflect ante. Surface reflect ante occurs at the interface between materials; usually the air-material interface and can be quite anisotropic. Subsurface reflectance is typically isotropic and occurs when incident light penetrates the surface and collides with inhomogeneities (such as paint pigments) suspended in the substrate. For rough matte surfaces, the observed reflectance is usually diffuse (Lambertian) and therefore nondirectional. The BRDF in this case is independent of both the incident and exitant directions and can be expressed as

$$\rho'_{bd}(\Theta_i, \Theta_r; \lambda) \equiv \rho'_{bd}(\lambda). \tag{3}$$

Perfectly smooth planar surfaces on the other hand behave as specular reflectors that reflect light in the critical mirror angle only. In this instance, the BRDF is highly dependent on the incoming and outgoing directions and can be expressed as a Dirac distribution.[18] Most practical surface mat erials are somewhere between the two extremes of diffuse and specular reflectors. Therefore, BRDF models typically represent the surface reflectance in terms of both a diffuse component and a specular component.

The aim of BRDF models, then, is to predict the reflectance behavior of surfaces under conditions of incident radiant energy striking the surface from various directions. One such model is the Beard-Maxwell BRDF model.

2,2.1. Beard- Maxwell BRDF

The Beard-Maxwell (B-M) BRDF model[19] was originally developed to characterize the reflectance properties of painted surfaces but has been used successfully in estimating the reflectance of other surfaces as well. Empirically based, this model has seven input parameters that are derived from a series of reflectance measurements collected under the controlled conditions of a laboratory setting.

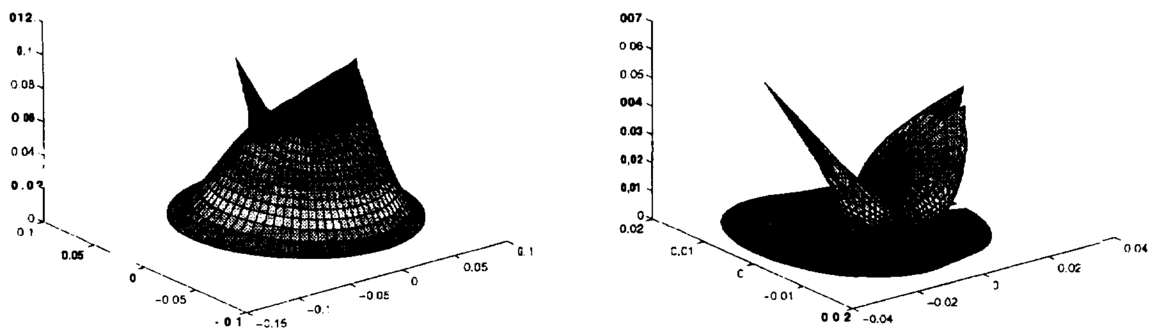
A functional description of the B-M BRDF model is specified by Equation 4

$$\rho'_{bd}(\Theta_i, \Theta_r) = \frac{R(\beta)}{R(0)} \frac{\rho_{fs} cos^2\theta_N}{cos\theta_i cos\theta_r} \left[\frac{1+\theta_N}{1+\frac{\theta_N}{\Omega} e^{-2\beta/\tau}} \right] + \rho_d + \frac{2\rho_v}{cos\theta_i + cos\theta_r} \tag{4}$$

where the first term in Equation 4 is the first surface reflectance attenuated by the shadowing and obscuration function enclosed in square brackets. The second term, ρ_d, is the diffuse (Lambertian) component and the third term is the component due to subsurface volumetric scattering.

3. Selection of Optimal Viewing Angles

In this section, the utility of BRDF models in the selection of optimal viewing angles for classification of remotely sensed images is explored. To motivate this discussion, consider the B-M BRDF of a galvanized steel rootop surface generated under the following conditions: the material is oriented in a horizontally flat position (θ_{matl} = 00), the solar position is specified by a zenith



(a) ρ'_{bd} (b) ρ'_s

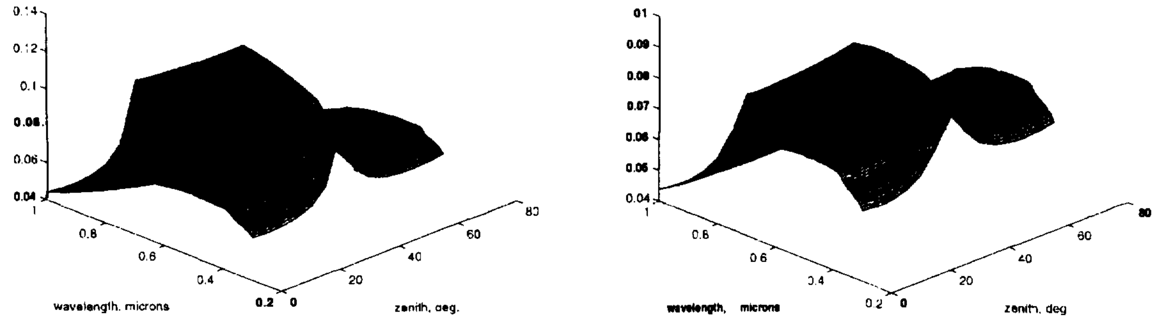
Figure 2. B-M BRDF: 0526stla (galvanized steel rooftop)

angle of θ_i = 30° and an azimuth of ϕ_i = 0°, with the BRDF evaluated at a spectral wavelength of $\lambda = 0.635\mu m$. Note that the observed BRDF seen in Figure 2(a) is dependent on viewing angles in both the azimuth and zenith directions due to forward scattering (sharp peak on the left) and backward scattering (peak on the right) effects. These effects are also seen in the surface component of the B-M BRDF shown in Figure 2(b) where once again the forward scattering peak is sharp and to the left with a significant backscattering lobe directed toward the solar position. Analagous to the variation in spectral signatures among material types, the spatial distribution of reflectance with angle gives rise to angular signatures specific to material type. Differences in angular signature can lead to an increased ability to distinguish between materials if identified and incorporated into a classification system. Our goal, then, is to identify regions of maximum difference between angular signatures through the use of the B-M BRDF model.

The B-M BRDF of several surfaces was generated with the the material positioned horizontally flat, the solar position given by a zenith of θ_i = 30° and an azimuth of ϕ_i = O“; wavelength ranged from 0.3- 1.0 μm in increments of 0.005 μm. Positioning of the sensor varied from ϕ_r = O“ to ϕ_r = 180° in the azimuthal direction at 5° increments and in the zenith from θ_r = 0° to θ_r = 65° every 5° corresponding to each azimuthal direction. As an illustration, B-M BRDF surfaces of a galvanized steel rooftop with the sensor positioned at an azimuth of ϕ_r = 180° and ϕ_r=175° are shown in Figures 3(a) and 3(b), respectively. Under similar conditions, B-M BRDF surfaces were generated for an aluminum painted low emissivit y green surface (0537 UUUPNT), a gray unweathered polyurethane paint on aircraft surface (0741 UUUPNT), and a concrete runway surface (0671 UUUCNC).

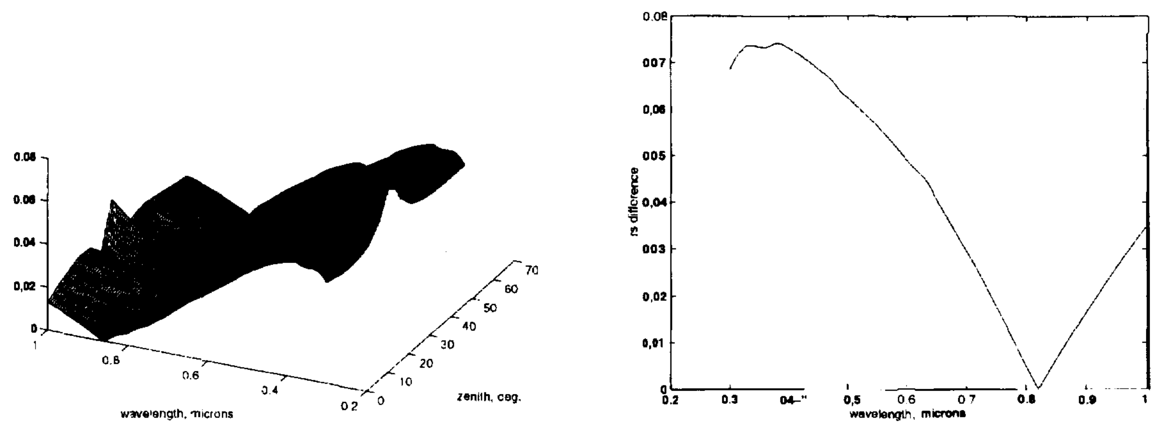
Pairwise separability was then computed on the basis of a pointwise root-squared (rs) difference between B-M BRDF surfaces of each material. Therefore, corresponding to each azimuth direction a root-squared difference surface was determined as shown in Figure 4(a). Inspection of the rs difference surface indicates that the maximum difference occurs in the zenith direction of θ_r = 30°. A cross-section of the surface along θ_r = 30° (Figure 4(b)) further shows that the maximum rs difference is achieved in the shorter wavelengths around $0.4\mu m$ and decreases as wavelength gets longer. A similar analysis in each azimuthal direction, ϕ_r = {O, 5,10, ...60,65}, resulted in the maximum difference between BRDF surfaces occuring at ϕ_r = 0°, θ_r = 30°, and in the wavelength range of 0.4 – $0.5\mu m$.

A separability analysis to determine the optimal viewing conditions for discrimination between



(a) $\phi_r = 180°$ (b) $\phi_r = 175°$

Figure 3. B-M BRDF Surface: 0526UUUSTLa (galvanized steel rooftop)



(a) $\phi_r = 180°$ (b) $\phi_r = 180°$, $\theta_r = 30°$

Figure 4. Pointwise Difference Surface: 0526 UUUSTLa (galvanized steel rooftop) and 0537UU-UPNT (green painted aluminum surface)

a concrete runway (0671 UUUCNC) and a painted aircraft surface (0741 UUUPNT) was also conducted. Results indicated that the maximum rs difference occurred in the primary backscatter direction, ϕ_r = 0° and optimal θ_r in the range 20° – 40°. Additional analysis indicated that the optimal wavelength range for discrimination, however, was in the longer wavelengths from approximately 0.8 μm to 1.0 pm.

4. Conclusions and Future Work

The utility of BRDF models for pairwise discrimination between surface materials through the selection of optimal view angles was explored. Employing a strategy of maximum rs difference between the respective BRDFs to indicate maximum separability, several cases were considered of which two were reported in this paper. In every case studied, the maximum rs difference occurred in the principal plane of the Sun with the BRDF due to backscatter effects yielding a slightly greater separability than the forward scattering direction.

It should be noted, however, that there were several limitations in this preliminary study: our analysis at this point was qualitative, atmospheric effects were completely ignored (we expect that these effects are not significant at low altitues but could be extremely important at high altitudes), and only one solar position was considered.

Future efforts will concentrate on quantifying the search for optimal viewing angles, incorporating atmospheric effects, and exploring the affect that solar position has on the separability y between surface BRDFs.

REFERENCES

1. P. H. Swain and S. M. Davis, *Remote Sensing: The Quantitative Approach*, McGraw-Hill, 1st. ed., 1978.
2. T. M. Lillesand and R. W. Kiefer, *Remote Sensing and Image Interpretation*, John Wiley & Sons, Inc., 3rd cd., 1994.
3. J. A. Richards, *Remote Sensing Digital Image Analysis: An Introduction,* Springer-Verlag, 1st. cd., 1986.
4. P. E. Ardanuy, D. Han, and V. V. Salomonson, "The moderate imaging spectrometer (modis) science and data system requirements," *IEEE Trans. on Geosci Remote Sensing* **GE-29**, pp. 75-88, Jan. 1991,
5. D. J. Diner and et al., "Misr: A multiangle imaging spectroradiometer for geophysical and climatological research from eos," *IEEE Trans. on Geosci Remote Sensing 27,* pp. 200-214, 1989.
6. J. R. Irons, K. J. Ranson, D. L. Williams, R. R. Irish, and F. G. Huegel, "An off-nadir-pointing imaging spectroradiometer for terrestrial ecosystem studies," *IEEE Trans. on Geosci Remote Sensing* **GE-29**, pp. 66–74, Jan. 1991.
7. National Photographic Interpretation Cent er, *Spectral Catalog Version 4.0,* Jan. 1993.
8. B. Holben and R. S. Fraser, "Red and near-infrared sensor response to off-nadir viewing," *Int. J. Remote Sens. 5,* pp. 145-160, 1984.
9. B. Holben and D. Kimes, "Directional reflectance response in avhrr red and near-ir bands for three cover types and varying atmospheric conditions," *Remote Sens. Environ. 32,* pp. 213-236, 1986.
10. R. D. Jackson, P. M. Teillet, P. N. Slater, G. Fedosejevs, M. F. Jasinski, J. K. Aase, and M, S. Moran, "Bidirectional measurements of surface reflectance for view angle corrections of oblique imagery," *Remote Sens. Environ. 32,* pp. 763–766, Nov. 1976.
11. G. M. Foody, "The effects of viewing geometry on image classification," *Int. J. Remote Sensing 9(12), pp. 1090–1915, 1988.*
12. D, S. Kimes and J. A. Kirchner, "Radiative transfer model for heterogeneous 3-d scenes," *Applied Optics 21, pp. 4119–4129,* Nov. 1982.
13. S. A. W. Gerstl and C. Simmer, "Radiation physics and modelling for off-nadir satellite-sensing of non-lambertian surfaces," *Remote Sens. Environ. 20,* pp. 1-29, 1986.
14. B. Pinty, M. M. Verstraete, and R. E. Dickinson, "A physical model for predicting bidirectional reflectances over bare soil," *Journal of Geophysical Research 27,* pp. 273–288, 1989.
15. K. E. Torrance and E. M. Sparrow, "Theory for off-specular reflection from roughened surfaces," *J. Opt. Sot, Am, 57,* pp. 1105-1114, Sept. 1967.
16. X. Li and S. A. H., "Geometrical-optical bidirectional reflectance modelling of a conifer forest canopy," *IEEE Trans. on Geosci* Remote Sensing **GE-24**, pp. 906-919, 1986.
17. C. L. Walthall, J. M. Norman, J. M. Welles, G. Campbell, and B. L. Blad, "Simple equation to approximate e the bidirectional reflectance from vegetative canopies and bare soil surfaces ," *Applied Optics* 24, pp. 383–387, Feb. 1985.
18. W. G. Rees, *Physical Principles of Remote Sensing,* Cambridge University Press, 1st. cd., 1990.
19. J. Beard and J. R. Maxwell, "Bidirectional reflect ante model validation and ut ilizat ion," Technical Report TR-73-303, AFAL, Oct. 1973.

,,, ,.,8

5/30/47

URC97130

Regularization of Atmospheric Temperature Retrieval Problems*

Miguel Vélez-Reyes   Ruben Galarza-Galarza
Tropical Center for Earth and Space Studies
Electrical and Computer Engineering Department
University of Puerto Rico, Mayagüez, PR 00681-5000
Tel. (787) 832-4040 Ext. 3086,   FAX (787) 831-7564
e-mail: mvelez@exodo.upr.clu.edu

1 Introduction

Passive remote sensing of the atmosphere is used to determine the atmospheric state. A radiometer measures microwave emissions from earth's atmosphere and surface. The radiance measured by the radiometer is proportional to the brightness temperature. This brightness temperature can be used to estimate atmospheric parameters such as temperature and water vapor content. These quantities are of primary importance for different applications in meteorology, oceanography, and geophysical sciences. Depending on the range in the electromagnetic spectrum being measured by the radiometer and the atmospheric quantities to be estimated, the retrieval or inverse problem of determining atmospheric parameters from brightness temperature might be linear or nonlinear. In most applications, the retrieval problem requires the inversion of a Fredholm integral equation of the first kind making this an ill-posed problem. The numerical solution of the retrieval problem requires the transformation of the continuous problem into a discrete problem. The ill-posedness of the continuous problem translates into ill-conditioning or ill-posednessof the discrete problem. Regularization methods are used to convert the ill-posed problem into a well-posed one.

In this paper, we present some results of our work in applying different regularization techniques to atmospheric temperature retrievals using brightness temperatures measured with the SSM/T- 1 sensor. Simulation results are presented which show the potential of these techniques to improve temperature retrievals. In particular, no statistical assumptions are needed and the algorithms were capable of correctly estimating the temperature profile corner at the tropopause independent of the initial guess.

2 Radiative Transfer Theory in the Microwave Region

Radiative transfer theory describes the intensity of radiation propagating in a general class of media that absorbs, emit, and scatter the radiation [5]. The radiative transfer equation for a plane-parallel atmosphere is given by

$$\cos\theta \frac{d\,I_\nu}{dz} = -\sigma(z)I_\nu + J_\nu(z) \tag{1}$$

where I_ν (z) is the instantaneous radiant intensity that flows at each point in the medium per unit area, per unit of solid angle, at a given frequency ν; a(z) is the extinction coefficient; and J is a source term. These last two quantities describe the loss/gain into the given direction. The angle θ is the direction angle with respect to the vertical axis z with θ = O when pointing upwards.

In the general case, scattering into and from other directions can lead to both gains and losses to the intensity and are taken care by the terms σ and J. For the microwave region, the scattering term is usually neglected [3]. If scattering is neglected, the only source term to consider is that due to local emission and the extinction coefficient reduces to the absorption coefficient σ_a. Assuming local thermodynamic equilibrium, each point can be characterized by a temperature T and from Kirchoff's law we get

$$J_\nu(z) = \sigma_a(z)B_\nu\ (\mathrm{T'(z)}) \tag{2}$$

*Research Parialy Sponsored by AFOSR Summer Faculty Research Program and NASA URC under grant NCCW-0088

where $B_\nu(T)$ *is* the Planck function:

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{\exp h\nu kT - 1} \tag{3}$$

where h = 6.625×10^{-34} Js is Planck's constant, c = 2.988×10^8 m/s is the speed of light, and $k = 1.381\times10^{-23}$ J/K. Equation (1) is a linear non-homogeneous first-order differential equation with solution

$$I_\nu(z) = I_\nu(z_o)e^{-\sec\theta\int_{z_o}^{z}\sigma_a(z)dz} + \sec\theta\int_{z_o}^{z}\sigma_a(z)e^{-\sec\theta\int_{\gamma}^{z}\sigma_a(\xi)d\xi}B_\nu(T(\gamma))\,d\gamma \tag{4}$$

where $I_\nu(z_o)$ is the boundary condition, In the microwave region of the spectrum

$$h\nu << kT$$

which results in (3) taking the form

$$B_\nu(T) = \frac{2\nu^2 kT}{c^2} = \frac{2kT}{\lambda^2} \tag{5}$$

where λ is the wavelength. This is known as the Raleigh-Jeans approximation. From this expression, it is clear that in the microwave region the energy emitted is proportional to the physical temperature T. Another commonly used result from this relation is to define a scaling of the intensity I_ν as follows

$$T_b = \frac{\lambda^2}{2k} I_\nu \tag{6}$$

The quantity $T_b(v)$ is called the ***brightness*** temperature which is commonly used in the microwave retrieval literature instead of I_ν. In terms of brightness temperature and using (5), (4) takes the form

$$T_b(z) = T_b(z_o)e^{-(\delta(z_o)-\delta(z))\sec\theta} + \sec\theta\int_{z_o}^{z}\sigma_a(\gamma)\,e^{-(\delta(\gamma)-\delta(z))\sec\theta}T(\gamma)d\gamma \tag{7}$$

where

$$\delta(z) = \int_z^\infty \sigma_a(\xi)\,d\xi \tag{8}$$

is the optical thickness and co represents the top of the atmosphere (TOA). Here the dependency of all these quantities in frequency v is not shown for convenience

For our purpose, it is of interest to solve this equation to obtain the brightness temperature that a satellite will measure at the top of the atmosphere when looking to the surface at an angle θ off nadir. This will correspond to z_o = O surface (sfc) and $z = \infty$ in (7). The boundary term T_b (0) is given by

$$T_b(0) = \epsilon T_s + (1 - \epsilon)T_d$$

where T_s is the surface temperature, T_d is the downwelling radiation reflected by the surface back towards the satellite, and ϵ is the surface emissivity. For the reflected component, it is assumed that the surface is a smooth, homogeneous, and isothermal so only the radiation in the specular direction θ is accounted for. In our reference coordinates, the propagation angle for the downwelling radiation is $\pi - \theta$. The downwelling radiation is obtained by solving the radiative transfer equation where integration is from the TOA to the surface with propagation angle $\pi - \theta$. This is taken care in (7) by setting $z_o = \infty$ and z = O. The boundary term in this case is given by the cosmic microwave background emission with T_c = 2.7 K. The resulting expression for the brightness temperature at the TOA at an angle θ off nadir

$$T_b = \left[\epsilon T_s + (1-\epsilon)T_c e^{-\delta(0)\sec\theta}\right] e^{-\delta(0)\sec\theta} + \int_0^\infty T(z)\left[1 - (1-\epsilon)e^{-2(\delta(0)-\delta(z))\sec\theta}\right] e^{-\delta(z)\sec\theta}dz \tag{9}$$

Simulation of this expression if all quantities were known is a simple matter. However, the computation of the optical thickness and the absorption parameter requires the use of databases containing information about the spectral characteristics of atmospheric constituents such as HITRAN. In our work, all optical depth computations were carried out using the FASE Radiative Transfer Code.

3 Atmospheric Remote Sensing

The satellite instrument measures radiance that arrives into its field of view. The radiance that arrives is the sum of the radiance emitted and reflected by the surface, emitted and reflected by the atmosphere, and that scattered by the atmosphere into the field of view of the instrument. The relative contribution of each component depends on the region of the spectra seen by the instrument. Remote sensing of the surface takes advantage of those regions of the spectrum where the atmosphere is transparent or nearly so. In the case of atmospheric remote sensing, the satellite sensor is looking at regions in the spectra where the atmosphere blocks the radiance emitted or reflected by the surface and therefore it receives that radiation that is emitted or reflected by the atmosphere. The interaction of electromagnetic waves with the atmosphere depends on the characteristics of the propagating wave (primarily its wavelength), the physical characteristics of the atmosphere and its constituents (pressure, temperature, density, absorbing gases, suspended particles). The mechanisms for interactions are: scattering, absorption, emission, and refraction. In regions where the atmospheric constituents characteristics are known or understood as in the 60 GHz oxygen and 183 GHz water vapor absorption lines, measurements of brightness temperature can be used to infer atmospheric properties of interest. The relation between atmospheric properties with brightness temperature is given by the radiative transfer equation (7). Therefore the problem of interest is to infer the atmospheric quantities of interest from measured brightness temperature by inversion of the radiative transfer equation.

4 Temperature Retrieval Problem

If the atmosphere strongly absorbs, most of the contribution to the measured brightness temperature will come from the atmosphere itself. In the case of the microwave region of the spectrum, scattering is negligible and the energy into the field of view of the sensor will come from atmospheric emission. Assuming that the satellite is looking at nadir (i.e. θ = O), and that the surface temperature T_s and emissivity ϵ are known, (9) can be rewritten as

$$\tilde{T}_b = \int_0^\infty \mathrm{T(z)}\ K(\nu, z)\, dz \tag{10}$$

where

$$\begin{aligned} K(\nu, z) &= \left[1 + (1-\epsilon)e^{-2(\delta(0)-\delta(z))}\right] e^{-\delta(z)} \\ \tilde{T}_b &= T_b - \epsilon T_s + (1 - \epsilon)\, T_c e^{-\delta(0)}\, e^{-\delta(0)} \end{aligned}$$

If the absorber is uniformly mixed with a known concentration, as 0_2, the quantity $K(\nu, \mathrm{z})$ is known and the temperature profile T(z) could be retrieved by inverting (10). The function $K(\nu, \mathrm{z})$ is called in the literature [3] the *weighting function*. From this point on, in our discussion we would not distinguish between T_b and $\tilde{T}_b$ in (10). 0_2 has several absorption lines around between 50 and 60 GHz. The Special Sensor Microwave Temperature-1 (SSM/T-1) sensor of the sensor suite of the DMSP satellite has 7 channels located in the 50 to 60 GHz range used for temperature retrievals. A summary of the SSM/T-1 sensor characteristics is given in [6]. In the 0_2 band, the shape of the weighting function is independent of the temperature making the inversion of (10) a linear inversion problem.

4.1 Problem Discretization

To numerically solve the temperature retrieval problem, first we discretize (10) by approximating the integral with a numerical integration formula. This results in the algebraic linear system of equation

$$\mathbf{T}_b = \mathbf{KT} + \mathrm{e} \tag{U}$$

where $\mathrm{T}_b \in \mathcal{R}^m$ and $\mathrm{T} \in \mathcal{R}^n$ are the brightness and atmospheric temperature vector; $\mathrm{K} \in \mathcal{R}^{m \times n}$ is the matrix of weighting functions; and e is an error term associated with measurement noise and the truncation error arising from the discretization of the integral equation. The number of measured brightness temperatures m is usually smaller than the vertical resolution or number of temperature levels n to estimate. In the case of the SSM/T-1 sensor, there are m = 7 channels and normally n $\geq$20 temperature levels. Therefore the resulting algebraic linear system of equations (11) is under constrained (i.e. there are more unknowns than equations).

4.2 Regularization of the Discrete Problem

The temperature retrieval problem is related to the solution of the linear system of equations (11). This problem has two major difficulties associated with it: (i) ill-conditioning due to the ill-posedness of the associated integral equation, and (ii) multiple solutions because of trying to estimate more temperature levels than measurements available. To overcome these difficulties, we will use the so called regularization methods. Regularization theory [1] transforms an ill-posed problem to a well-posed one, using a *priori* knowledge on the nature of the solution. Depending on the prior information, regularization techniques can be classified into two major groups: statistical and deterministic.

4.2.1 Statistical Regularization

In statistical regularization, prior statistical information is used to regularize the temperature retrieval problem. Our prior information in this case is the prior distribution of the temperature profiles $p_T(T)$ and the conditional distribution $p_{T_b/T}(T_b)$. According to Bayesian estimation theory [4], the best estimator T based on the brightness temperature observation $\mathbf{T}_b$ of the temperature profile T in a mean square sense is the conditional mean

$$\mathrm{T} = E\,(\mathbf{T}/\mathbf{T}_b) \tag{12}$$

We will refer to this estimator as the minimum mean square estimator (MMSE). The analytical determination of this function might be a very difficult task. In many instances, the estimator is constrained to be linear which results in the Linear Minimum Mean Squares Estimator (LMMSE) [4]

$$\mathbf{T} = \mathbf{T} + \Lambda_{T,T_b} \Lambda^{-1}_{T_b,T_b} (\mathbf{T}_b - \bar{\mathbf{T}}_b) \tag{13}$$

where T is the a priori mean of T, $\bar{\mathbf{T}}_b$ is the mean of $\mathbf{T}_b$ given T, Λ_{T,T_b} is the cross covariance between T and $\mathbf{T}_b$, and Λ_{T_b,T_b} is the conditional covariance of the brightness temperature. The LMMSE is easy to construct, since only the first and second order statistics are needed rather than their complete probability densities. Also, if T and $\mathbf{T}_b$ are jointly Gaussian, the LMMSE is the optimal Bayesian MMSE.

4.2.2 Deterministic Methods

In this section, we will look at two regularization methods for ill-posed linear algebraic systems of equations: Tikhonov regularization, and discrepancy principle regularization. Other methods are discussed in [2]. Computation of the regularized solution was done using the MATLAB Regularization Toolbox presented in [2].

Tikhonov's Regularization

One way to regularize (11) is computing T as the solution to the optimization problem

$$\hat{\mathbf{T}}_\lambda = \arg \min_{T \in \mathcal{R}^n} \left[\|\mathbf{KT} - \mathbf{T}_b\|^2 + \lambda^2 \, \|\mathbf{L}\,(\mathbf{T} - \mathbf{T}_o)\,\|^2 \right] \tag{14}$$

where TO is a prior temperature profile estimate, $\|\cdot\|$ is the 2-norm, and λ is the regularization parameter.

A key issue in this method is the selection of the regularization parameter λ. The value used in the simulation results presented here was based on the L-curve method described in [2]. The optimal value of λ balances the prediction error $\|\mathbf{KT} - \mathbf{T}_b\|$ with the regularization error $\|\mathrm{L}(\mathrm{T} - \mathrm{TO})\|$.

Discrepancy Method

Another possibility to regulate the temperature retrieval problem is by computing T as the solution to the quadratically constrained linear least squares problem

$$\hat{\mathbf{T}}_\alpha = \arg \min_{T \in \mathcal{R}^n} \|\mathbf{L}\,(\mathbf{T} - \mathbf{T}_o)\|^2 \tag{15}$$

$$\text{subject to} \quad \|\mathbf{KT} - \mathbf{T}_b\|^2 \le \alpha^2$$

where α plays the role of a regularization parameter. The solution to this problem can be made identical to $\hat{\mathbf{T}}_\lambda$ for a suitably chosen α [2]. We prefer to select α based on the measurement noise norm $\|\mathrm{e}\|$.

5 Simulation Experiments

In this section, we present some simulation results that illustrate the use of deterministic methods to regularize the temperature retrieval problem. For the simulation experiments, a 98 layer atmosphere based on the US Standard Atmosphere model was used. The optical thickness for each layer was computed using the FASE code. The radiative transfer equation was numerically integrated using the trapezoidal rule. The implementation of Tikhonov's and the discrepancy algorithm available in the Matlab^{TM} Regularization Toolbox were used to compute the temperature retrievals. All computations with the exception of the optical thickness were done under the MATLAB^{TM} environment. In our simulations, the surface emissivit y ϵ was set to 0.9 and the surface temperature T_s set to 288.2 degrees Kelvin. Figures 1 and 2 show the results of applying Tikhonov's regularization to the temperature retrieval problem with perfect measurements (no noise). The regularization parameter λ was set 0.0044. The solid line is the retrieved temperature profile, the dashed line is the actual temperature profile, and the dash-dot line is the initial guess fed to the algorithm. We can see from this simulations that the algorithm was capable of estimating temperature up to 40 km. Beyond 40 km, the resulting estimate was identical to the initial guess. An important result is how the retrieval algorithm is capable of determining the height of tropopause corner when fed with initial guesses that have that corner at heights far from the actual height as shown in Figure 2. The importance of determining this peak comes from the fact that most important weather features are located at this region of the atmosphere. Also, the location of the tropopause peak serves as a figure of merit in evaluating the performance of temperature retrieval algorithms. We are not showing results for the algorithm based on the discrepancy principle regularization since they were similar to those of Tikhonov regularization.

Figures 3 and 4 show the performance of Tikhonov regularization under the presence of noisy data. The noise vector added to the measured brightness temperature has a normal distribution with zero mean and unit variance. The regularization parameter for this case was $\lambda = 0{,}016$. Notice that the noise causes the retrieved profile to be noisier with a maximum error in the first 20 km of 10 degrees Kelvin. Quite high compared to some retrieval methods that claim accuracies of 0.5 degrees Kelvin. However, the algorithm is still capable of determining the height of the tropopause.

Figures 5 and 6 show the results for the retrievals in the noisy case computed using discrepancy principle regularization. The regularization parameter for this case was set at $\alpha = 3$. That value corresponds to three standard deviations of the noise distribution. The resulting retrievals are smoother than those retrievals from Tikhonov regularization. The maximum error is in the neighborhood of 3 to 4 degrees Kelvin in the tropopause. The estimation of the location of the tropopause peak is also improved.

6 Conclusions and Final Comments

This paper presents some preliminary work in the application of regularization techniques to the linear problem of atmospheric temperature retrievals from microwave radiometry. The proposed techniques were evaluated using simulated data. We used algorithms implemented in the MATLAB Regularization Toolbox developed by [2]. The results obtained were quite encouraging. In particular, being able to estimate the location of the tropopause corner of the temperature profile even with bad initial guess and noisy data is a result not previously observed with other algorithms.

References

[1] Engl, H. W., M. Hanke, and A. Neubauer, *Regularization of Inverse Problems.* Kluwer Academic Publishers, 1996.

[2] Hansen, P.C. *Rank-Deficient and* Discrete *Ill-Posed Problem..*, Ph.D. Thesis, 'Technical University of Denmark, 1996.

[3] Janssen, M.A. *Atmospheric Remote* Sensing *by* Microwave Radiometry, John Wiley & Sons, 1993.

[4] Kazakos, D. and P. Papantoni-Kazakos. *Detection and Estimation,* Computer Science Press, 1990.

[5] Lenoble, J. *Atmospheric Radiative Transfer.* A. Deepack publishing, 1993.

[6] Neu, T.J. "Defense meteorological satellite program microwave radiometer processing at air force global weather central," In *Proceedings of the First NMC/NESDIS/DOD Conference on DMSP Retrieval Products,* April 14-15, 1992. Phillips Laboratory, Report PL-TR-92-2191.

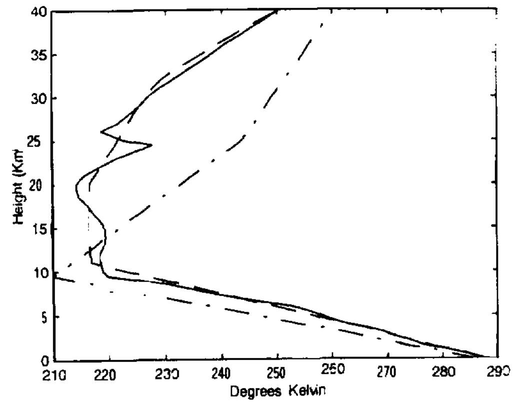


Figure 1: Tikhonov Regulatization with 10 km Tropopause.

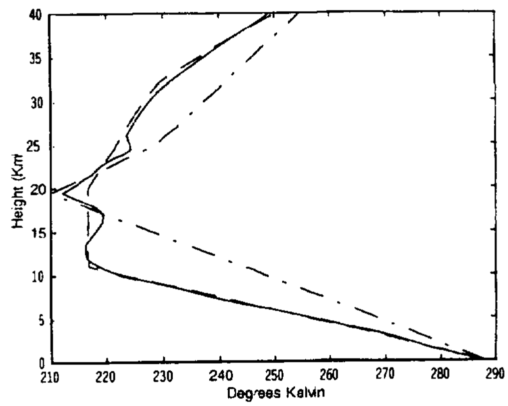


Figure 2: Tikhonov regularization with 20 km tropopause initial guess: noise free.

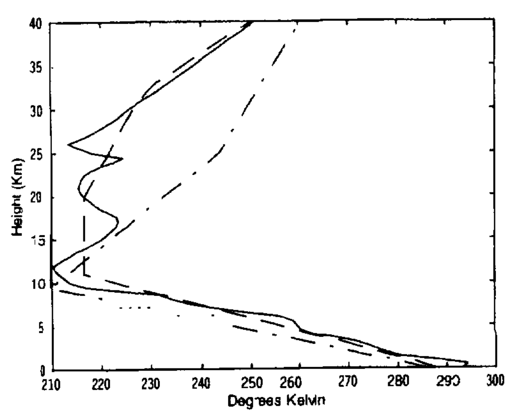


Figure 3: Tikhonov regularization with 10 km tropopause initial guess: noisy case.

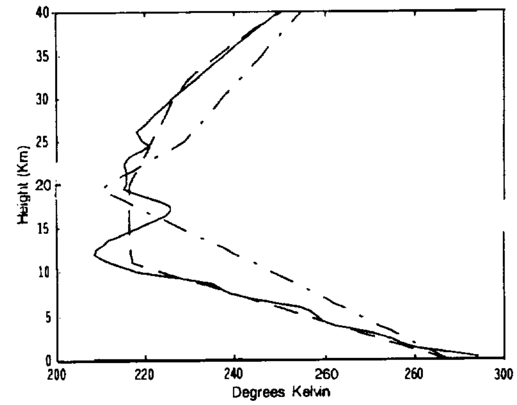


Figure 4: Tikhonov regularization with 20 km tropopause initial guess: noisy case.

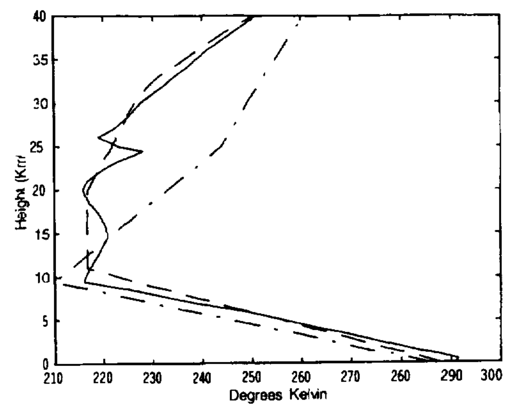


Figure 5: Discrepancy regularization with 10 km tropopause initial guess: noisy case.

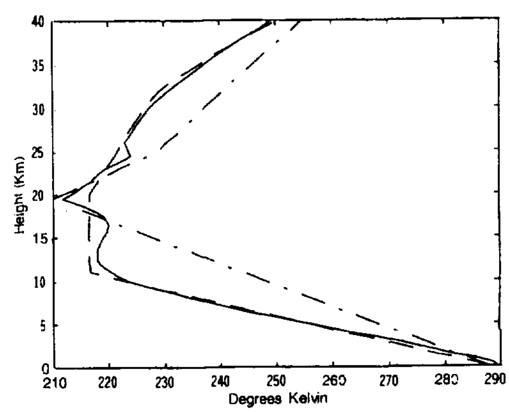


Figure 6: Discrepancy regularization with 20 km tropopause initial guess: noisy case.

URC97131

P. 771

Platform-Independent Teleoperation

Luis R. Villalta
Department of Engineering
Inter American University of Puerto Rico -- Bayamón Campus
Bayamón, PR 00959, USA
http ://www.bayamon. inter.edu

1 Introduction

Teleoperation and telerobotics have been for the most part platform-dependent with systems designed to work with specific equipment. Teleoperation here means low-level control of a robot, while telerobotics means control of a semi-autonomous robot, Platform-independence is used to denote systems that can work with different types of computers and controllers, and multiple operating systems. The main idea for this application is that the slave system can be controlled from different kinds of master systems. If one designs a system that is platform-independent above everything else, it is necessary to let go of graphics tools, object oriented development tools, and any other machine-specific utilities. The alternative is to find a "neutral" development environment that provides platform independence.

There are various areas where a specific platform may be needed depending on the kind of teleoperation application, Platform dependent capabilities provide sometimes indispensable information to the operator [3, 4], since there is often no visual feedback, and obviously no tactile feedback. These capabilities include graphical displays, video, and still photos. It is important to find graphical, video and picture capabilities that can be used under different platforms. A precisely developed geometric model of the mission, showing approximate current positions is often the only reliable source of information. Video and still photos are also commonly used, but not always available [3]. There is also the issue of security of the actual connection between master and slave, these connections can also rely on platform-dependent solutions. Here the tradeoff is between the level of security, and access flexibility.

The purpose of this project was to determine requirements for a general framework that would allow platform-independent teleoperation and telerobotics through the Web, thus providing a straightforward way of publicizing new algorithms, paradigms, and procedures, while maintaining high security and privacy. The goal was also to identify existing technologies that would allow this without a need for custom development, and provide reusability and straightforward adaptability into different types of applications. The framework should also allow non-experts to develop.

This paper describes the general characteristics and feasibility of such a framework, providing examples of current technology. It does not tout any specific software, but yet describes current capabilities available on the Internet. All techniques described can be implemented using existing paradigms that dynamically generate HTML pages on the server. The framework itself is described on section 6.

2 Data Communication

The Internet is used for data communication, although this framework could still be used on an intranet. Some notable attempts at making teleoperation feasible through the Web are applications like Project Mercury and the Telegarden at the University of Southern California (*http://www.usc.edu/Showcase/*), the Telerobot at the University of Western Australia (*http://telerobot.mech.uwa.edu.au/*), and Xavier at Carnegie Mellon University (*http://www.cs.cmu.edu/~Xavier/*). These sites have shown the feasibility of remote control through the Internet, and specifically through the World Wide Web.

There are implementation, security and privacy problems when connecting expensive equipment to the Internet. The problem caused by inexperienced operators controlling the equipment has been dealt with algorithms that block certain movements that may cause harm to the hardware, but there are still problems in maintaining platform-independent control while at the saline time protecting confidential, and maybe patentable, intellectual property.

There is also the problem of limited and varying bandwidth. This problem is also found in applications where the sequence of tasks is of extreme importance, and others where there are natural limitations such as space or undersea teleoperation. Teleprogramming is one of the attempts to shield the operator from these time delays product of a limited bandwidth [1, 2, 3, 4], though it is not the only attempt [5]. It was decided early onto consider the feasibility y of implementing teleprogramming through the World Wide Web as part of this study, where teleprogramming means an interface that allows teleoperation using high-level commands while shielding the operators from delays [2, 3, 4]. (See section 5.)

3 Security

The dilemma is how to protect was is being made public, since the World Wide Web is a public medium. The specific applications that were considered, although these are only examples, are a home page to publicly display the capabilities of a system, and an interface to allow flexible remote access to authorized personnel. The framework is effective on applications where we can allow the public to interact with a system, but want to maintain control over the code itself. It can make public the operation, while hiding the internal programming of an algorithm.

The first application considered was secure remote access. One approach is to set up the web server itself to allow or deny connections based on hostname, or on password protection. While this might be sufficient for some, less critical applications, security deficiencies of some software are public knowledge on the Internet. The alternative may be to code a local version of password protection, using an active page generation scheme on the web server, rather than working with submitted forms. This provides a second layer of security, while at the same time hiding the inner workings of the site.

A problem with forms is that in some implementations the user can just look at the source code of the web page and glean important information on how the server, or the application, works. By using a dynamic page generation scheme, the user only sees standard HTML and text when viewing the source of an HTML page. Dynamic page generation is an approach to generating HTML where "static" HTML pages are generated using programs running on the server side. These pages are standard HTML, they can be read by any web browser, yet graphics and even text can vary via user interaction. All internal variables are hidden, as well as the format used to submit information to the server. As the user tries to access the site, the system presents fields for username and password. The responses given are then compared with a file. This file could be on the server itself, or on some other remote machine connected to the server via a UNIX port. These features not only provide greater security, but also enhance performance, since multiple processes can be dynamically generated on different machines to support higher demand. (See section 4 and figure 1.)

The second application considered was the implementation of a public display for an algorithm or technique. One of the main concerns here is the way that multiple accesses are handled by web servers. One problem would be two or more remote operators trying to move the hardware on different directions. This can be handled by teleprogramming, but will also need to handle state variables for every user, since pages from different users can be served in alternating fashion. Dynamic page generation software, available on the Internet, provides a solution to this state problem, because it is normally encountered on "Web shopping" applications. If there are multiple users connecting to a server, and ordering different merchandise, the server sorts out who "owns" a specific dynamic page based on this state information. This solution will appear only

to apply to graphical simulations and not to actual teleoperation, since it is customary that only one operator at a time control a robot, but by implementing teleprogramming the user can be actually generating the commands that then will be executed as soon as possible.

There could be particular applications where the goal is to show how swiftly a task can be performed. This can require giving the operator complete control for a specific length of time. The development system allows for this, using username and passwords, while maintaining a list of current connections. The system can be overruled, adding functions implementing priorities or other restrictions.

One other issue related to security was the use of the so called "web languages" like Java, for example. While these languages introduce a great deal of flexibility to the development of a system, there is the matter of sending code through the Internet that can then be reverse-engineered. For normal applications it is secure enough to use platform-independent Java interfaces to communicate with the server. Java, and any other "web language" would be used only for graphical interfaces and not to implement algorithms on the client, since this would amount to "giving away" the algorithm. (See section 6.) Using the dynamic page software, some Java code can also be executed *on* the server side.

4 Performance

The two situations considered for performance were the response time of the server to a particular request, and the ability of the system to emulate platform-dependent teleoperation and telerobotics. The first one of these deals with the actual performance of the Internet connection under normal conditions, while the second concerns the suitability of the general framework to implement graphical and other types of interfaces of the kind needed by teleprogramming and other applications.

The first condition was satisfied because the dynamic page generation scheme tested provided support for the generation of multiple parallel processes on different machines communicant ing with the server via UNIX ports. Once a dynamic application is created, it can be installed on different machines on the network even if these do not have a web server daemon running. Once the load on the server exceeds some specified limits, the system can invoke new instances of the application on different machines. A registered version of the software used can communicate with normal personal computers running office applications even allowing, for example, a spreadsheet to receive data from, and output calculations to, the Web.

The second condition can be satisfied with a proper design of the interface, using technologies such as the Virtual Reality Modeling Language (VRML) to generate graphical models. The VRML 2.0 standard provides the basis for interaction and animation, so there is a possibility that it could be used in a suitable interface, (See section 5.)

5 Interface

The interface tends to be one of the most important parts of a teleoperation system. Interfacing issues included from the actual networking between machines, up to graphical interfaces developed in Java, and modelling using VRML 2.0. Teleprogramming adds its own set of requirements to the graphical interface, since it needs three-dimensional graphical models.

A suitable interface can be built using Java, or other "web language" to provide platform-independent graphical interfaces, and VRML to provide 3D modelling capabilities. The dynamic page generation software on the server can setup different views of the model. The networking issue can be handled using standard communications protocols.

The interface developed for the feasibility study included Java controls that communicated with the server-side software, and straight HTML text forms. The server-side software provides support for all other types of interface just discussed.

6 General Framework

To really get a feel for what requirements would be like in an actual application, in order to build a more realistic picture of the system, a sample application was created. This sample application included as many features as possible of those that were needed in a teleoperation implementation, with or without teleprogramming.

Feasibility was determined using NeXT Software's WebObjects, which is available for Windows NT, SUN Solaris, and NEXTSTEP/OPENSTEP. The scripting language included with the freeware version of the package provides functionality to develop scripted applications that can run on different machines and are dynamical y generated (*http://www.next.com/WebObjects/*). The final version of this sample application has a text based interface demonstration, and a small Java control that can run on the client, a Java-enabled browser, and transmit data to the WebObjects application on the server, or another machine connected to the server.

The general framework provides support for all content types available now on the World Wide Web. The general organization is as shown on figure 1; the server generates HTML pages based on information provided by the user, and applications or scripts running on machines on the local network. The robot controller is one of those machines. Interfaces can be generated in any of the" web languages" now available, to provide a GUI look and feel to the operator.

A minimum set of characteristics needed to develop applications on the Web that support telerobotics, and produce reusable, encapsulated components would be as follows:

- Object Oriented Development environment, emphasis on Open platforms
- Platform-independent graphical user interface (GUI), using Java or some other "web language"
- Support for graphical "CAD style" modelling using VRML (*Virtual Reality Modelling Language*), and VRML 2.0
- Standard HTML (Hypertext *Markup Language*) output that is easily read by different kinds of browsers, even text-based ones if need be
- Web server independence - the application can be served from any http daemon
- Support low-bandwidth connections
- Provide high security of the actual code of algorithms, supporting the Secure Socket Layer (SSL) or other security protocols on the web, as well as allowing the developers to devise their own security schemes locally

7 Results and Conclusions

The trial application consisted of a couple of linked HTML pages that were dynamically generated by

WebObjects, and were served from a machine connected to the same network as the web server where the **http** request was sent. The first page included some text, and name and password fields used to control access to the pages following. The next page included the Java GUI controls that allowed the client to send data to the server, as well as some text based controls. These would be suitable for supervisory control.

What is really different about the approach, is that it allows the developer to generate standard HTML code dynamical ly, that can then let the user know what is happening interactively. An example of this is how counters are handled at some sites that use this dynamic page generati on scheme, such as the download page for WebObjects (*http://www. next.com/WebObjects/*). Normall y these counters are implemented only with graphics, so that they look like odometers to graphical browsers, but just present the static alternate text "[Picture]", or something similar, to a text browser like Lynx. If one browses that dynamic page with Lynx,

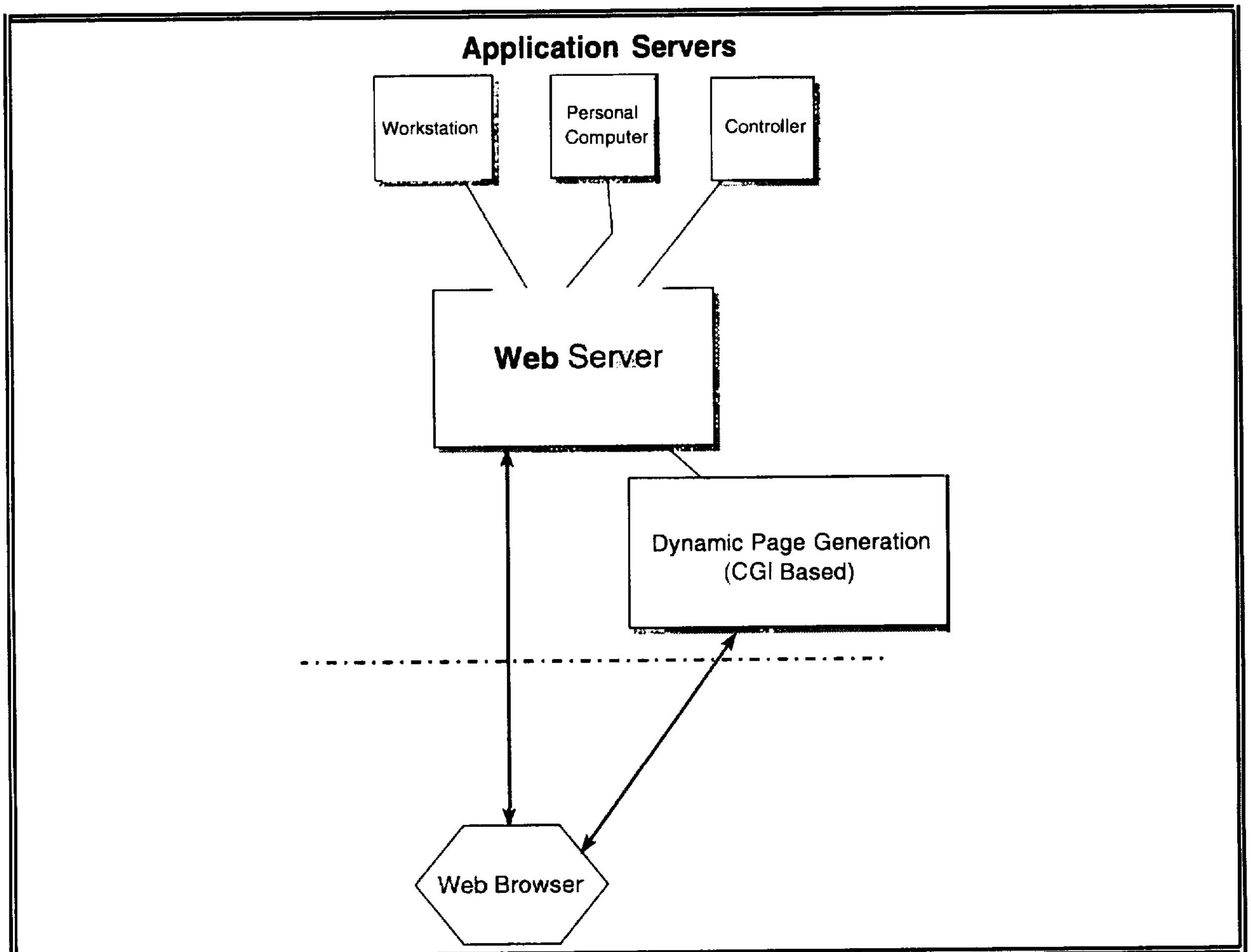


Figure 1 This is a graphical representation of the general framework. Note that there could be content that proceeds directly from the server, while other is filtered through the dynamic page generation CGI.

the numbers change in the alternate text, so one can see the actual count on a text based browser. This would even allow the creation of alternate character based graphics to show the position of a manipulator on different axes on a text browser.

This same flexibility would allow the server to send the user a dynamically chosen VRML 2.0 file that would show the movement of that manipulator, given the commands issued. This could then provide a way of

implementing a teleprogramming interface through the web.

Since these dynamic page generation platforms have been developed for Web commerce, there is a straightforward application of security paradigms, but they also allow interaction with databases, even on mainframe computers. This capability can also help in developing the telerobotics applications, since now objects can be stored on databases, and not only text data.

It can be concluded that an implementation that is platform independent should rely on open standards. Object standards such as CORBA, and even OLE, can also help in the straightforward development of these kinds of applications. It is also fair to say that, even though the current state of the Web, and the Internet in general, is not conducive to port real-time applications to it yet, there are some applications, such as space telerobotics and Untethered Underwater Vehicles (UUV' s), that are not real-time by their physical limitations, and these can be ported now.

8 Acknowledgements

The author would like to acknowledge the support given by the Center For Autonomous Control Engineering, University of New Mexico. All work in this project was conducted at their facilities. A special mention is due to Dr. Francisco Tomei, whose initiative prompted this project. Dr. Tomei's vision is the reason why I was given the opportunity to travel to Albuquerque, and get to know the fields of robotics and autonomous controls in a new light. His help on polishing up this paper is greatly appreciated.

9 References

[1] C. Sayers, M. Stein, A. Lai, R. Paul. Teleprogramming to Perform Sophisticated Underwater Manipulative Tasks Using Acoustic Communications. In *Proceedings IEEE Oceans,* pages 168-173, September 1994.

[2] Janez Funda, Thomas S. Lindsay, and Richard P. Paul. Teleprogramming: Toward delay-invariant remote manipulation. *Presence,* 1(1):29-44, Winter 1992.

[3] Craig Sayers, Angela Lai, and Richard Paul. Visual imagery for subsea teleprogramming. In *IEEE International Conference on Robotics and Automation,* pages 1567-1572, May 1995.

[4] Craig Sayers and Richard Paul. An operator interface for teleprogramming employing synthetic fixtures. *Presence,* 3(4):309-320, *1994.*

[5] Thomas Sheridan. Space teleoperation through time delay: review and prognosis. *IEEE Trans. on Robotics and Automation,* 9(5):592-606, October 1993.

[6] JyiShane Liu. Collective Problem Solving through Coordination in a Society of Reactive Agents. Carnegie Mellon University, Robotics Institute Technical Report CMU-RI-TR-94-23, June 1994.

[7] Richard M. Voyles, Jr. and Pradeep K. Khosla. Multi-Agent Perception for Human/Robot Interaction: A Framework for Intuitive Trajectory Modification. Carnegie Mellon Universit y, Robotics Institute Technical Report CMU-RI-TR-94-33, September 1994.

S131/27

FORMATION MECHANISM OF METAL COLLOIDS IN OXIDE GLASSES: SILVER IN ION-EXCHANGED SODA-LIME GLASSES

Paul W. Wang
Department of Physics and Materials Research Institute,
University of Texas at El Paso, El Paso, Texas 79968

The structural and compositional changes of the soda lime glasses during the formation of the silver colloids were analyzed by the X-ray Photoelectron Spectroscopy (XPS) in order to examine the silver colloid formation mechanism. The *in situ* behavior of silver and SiO_2 networks on the surfaces of silver ion-exchanged soda-lime glasses during heating and cooling processes in ultra-high vacuum was monitored. The results showed that silver diffuses toward the surface, precipitates, and crystallizes during heating and the total silver surface concentration is slowly increased during cooling. The concentration changes and binding energy shifts of oxidized and neutral Ag, a new non-bridging oxygen species (NBO*), and a new silicon species (Si[a]) were applied to deduce a disappropriation reaction mechanism of Ag^+ on the surface during annealing. The SiO_2 network is modified at temperatures below 350 "C to accommodate more silver on the surface and to balance the extra charge carried by the Ag^+. That the SiO_2 network polymerizes during annealing was deduced from the results of the higher binding energies of Si 2p and O 1s after annealing. This observation suggests that the reduction of the Gibbs free energies and the relaxation of tensile stress result in the formation of the silver colloids under thermal annealing,

I. INTRODUCTION

Recently, nonlinear optical properties of glasses introduced by embedded ultrafine particles/clusters in glass matrices were extensively investigated [1]. This nonlinearity depends on the size, density and distribution of the nanostructure particles [2,3]. The controlled-size nanophase particles were usually formed by ion implantation [4-6], ion exchange [7,8], and sol-gel [9] techniques and subsequent annealing. Even though the structure of the nanoparticles in the glass was studied by many different techniques such as electron microscopy [9-1 1], atomic force spectroscopy [1 1], x-ray absorption fine structure [12,13], and FTIR [14], no one has done, to the best of our knowledge, an *in situ* study of the aggregation of silver during heating and cooling by X-ray Photoelectron Spectroscopy (XPS) in a UHV environment.

In this paper, we report on the silver transport properties and corresponding structural modifications in silver ion-exchanged soda-lime glass in order to understand the silver colloid formation mechanism during annealing. Since the ion exchange process in soda-lime glass has proved to be one of the simplest ways to embed the metal ions in the glass networks [15], the silver/soda-lime system was chosen to investigate the thermal behavior of silver on surfaces even though potassium, iron, calcium or other cations inside the glass may affect diffusion of the silver [16,17]. Further studies of the influence of cations in silver ion-exchanged glass during heat treatment are under way. Because XPS can reveal the chemical states of the diffusing species of silver and the dynamic modification of the SiO_2 network at various temperatures, it is a powerful tool for studying the *in situ* thermal behavior of silver on the surfaces of the silver ion-exchanged glasses.

IL EXPERIMENTAL

Silver ion-exchanged soda-lime glasses, one commercially available non-ion-exchanged soda-lime glass (72.6% SiO_2, 15.4% Na_2O, 6.3% CaO, 3.87% MgO, 1.63 Al_2O_3 and 0.2% Fe_2O_3), a pure SiO_2 glass sample (Spectrosil), and pure silver metal were investigated by XPS. The base pressure of the chamber during experiments was maintained at $1 \sim 3 \times 10^{-9}$ Torr. The silver ion exchange process and the experimental details were described previously [18]. The fluctuation in sample temperature is $\pm$ 5 "C and the uncertainty is $\pm$ 0.5% in relative concentration. The nearest neighbors of Ag^+ inside the glass network are oxygens [19] and therefore changes in oxidation state of the silver result in changes in the oxygens. It is known that bridging and non-bridging oxygens (BOs and NBOs) can be used to monitor structure changes in SiO_2 networks [20-22], To distinguish non-bridging oxygens from bridging oxygens the deconvolution of the asymmetric O 1s peak was carried out, where the FWHM of the bridging oxygen peak was fixed. The binding energy and FWHM of the non-bridging

oxygen component of the O 1s peak were allowed to adjust since the non-bridging oxygen concentration is expected to vary during the diffusion of silver toward the surface. Thus the overall best fit for the O 1s peak was obtained.

The binding energy of the core level photoelectron of an element was conventionally calibrated by assuming that the binding energy of the surface carbon 1s photoelectron is located at 284.6 eV [23]. The data obtained from these ion-exchanged samples were compared to those of pure silver metal, SiO_2 glass, and non-ion-exchanged soda-lime glass. The deconvoluted curves obtained from the fitting program and the binding energy of an element in the XPS spectrum were used to deduce the chemical bonding environment of the element. The surface diffusion species of silver, and the consequent dynamic modifications of silicon and oxygen on the surface were also deduced.

III. RESULTS

A. **Silver surface diffusion**

In order to investigate the chemical states of silver involved during annealing, pure silver with native oxide on it was studied by 2.8 keV argon ion sputtering. By comparing with a standard XPS spectrum of Ag 3d and noting the existence of a O1s signal, the XPS spectrum of Ag 3d obtained before sputtering was assigned to oxidized silver (Ag^+) and the one obtained after sputtering to neutral silver (Ago). The XPS spectra of Ag 3d from the ion-exchanged glasses were deconvoluted into oxidized states and neutral states by using a gaussian curve fitting program and using the binding energies of the two silver states obtained from post-sputtering results as guides. Fig. 1 illustrates the deconvolution of the Ag3d5/2 spectrum at a sample temperature 350 ‘C.

An *in situ* diffusion of silver toward the surface under heat treatment was monitored at sample temperatures from 20 °C to 450 *C at increments of 50 ‘C. Fig. 2 (a) to (f) show the XPS spectra scanning the photoelectron binding energy from 379 to 363 eV for a sample of ion-exchanged glass at various temperatures. As shown in Fig. 2, a steady growth in the intensities of Ag 3d signals strongly indicates thermal surface diffusion of silver. Also, it can be seen that the higher binding energy component corresponding to neutral silver in either the Ag 3d3/2 or 3d5/2 signal grows when the sample temperature increases. This clearly shows that the neutral silver on the surface gradually increases and becomes the dominant state during heat treatment. In other words, more Ag-O bonds break and more neutral silver is formed in the surface region during the heat treatment.

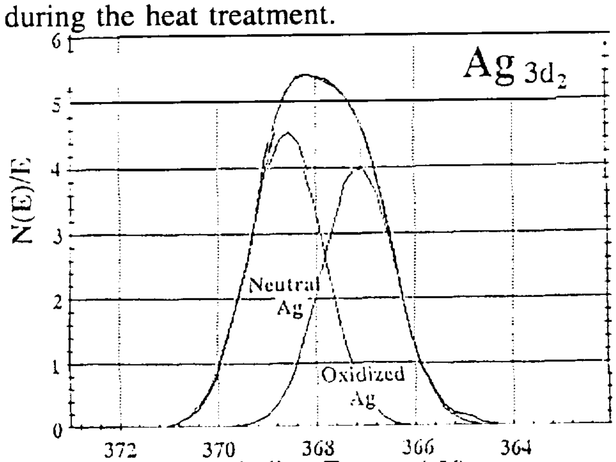


Figure 1. Deconvoluted Ag $3d_{5/2}$ spectrum of an ion-exchanged glass at 350 ‘C. The high binding energy (BE) component is neutral Ag and the low BE component is oxidized Ag.

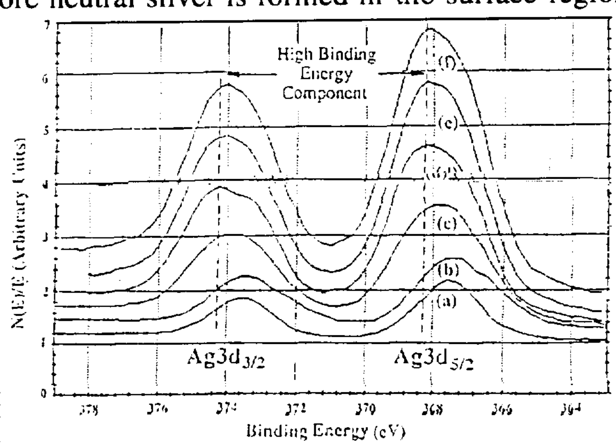


Figure 2. Ag $3d_{3/2}$ and $3d_{5/2}$ signals taken during surface diffusion on ion-exchanged glass at various temperatures, (a) 20 ‘C, (b) 100 ‘C, (c) 200 ‘C, (d) 300°C, (e) 400 ‘C, and (f) 450 ‘C.

The relative Ag atomic concentrations including total, oxidized, and neutral Ag during heating and cooling were calculated and plotted in Fig. 3, A three-step growth in total Ag is clearly seen during heating. During cooling the total Ag increases at a very slow rate but the interesting features are the relative concentration changes in oxidized and neutral Ag. The oxidized Ag increases rapidly between 450 ‘C and 220 °C and slows below 220 ‘C. In contrast,

an opposite trend was observed for neutral Ag, it decreases rapidly between 450 ‘C and 220 ‘C and then slows down below 220 °C.

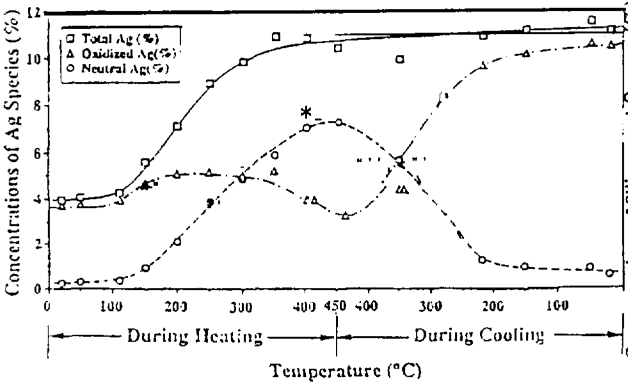


Figure 3. The relative surface concentrations of total Ag, oxidized Ag and neutral Ag during heating and cooling in ion-exchange glass.

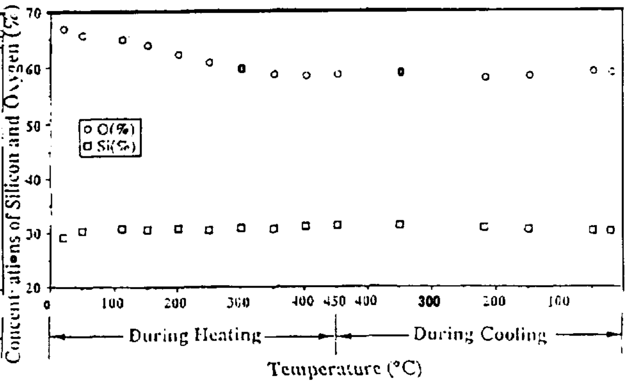


Figure 4. The relative concentration changes in Si and O during annealing.

B. **Changes of silicon and oxygen during annealing**

The relative atomic concentration changes in silicon and oxygen during annealing are shown in Fig. 4. It is seen that the changes in oxygen concentration with sample temperature show a decreasing pattern. The decrease rate slows down above 350 °C during heating and the concentration remains constant during cooling. There is no change in silicon concentration during annealing but two chemical species of Si were observed at temperature ranges from 110 to 450 °C and from 450 to 350 °C. A typical Si 2p spectrum with two components at sample temperature 300 ‘C is shown in Fig. 5. The high BE component belongs to Si atoms in a normal silicon-ox ygen network, i.e., silicon atoms surrounded by four oxygen atoms which include bridging or non-bridging oxygens [24]. and the low BE component is either caused by the normal network Si atoms each accompanied by one extra non-bridging oxygen [25], or by four-coordinated Si atoms each bonded to one negative doubly charged oxygen atom as discussed below. In order to distinguish these two types of Si atoms, we refer to the component with high BE as Si[n](normal) and to the component with low BE as Si[a](anomalous). Figure 6 shows the relative concentration of Si[a] in the whole Si 2p signal at various temperatures. It has a growth curve similar to the total Ag concentration in Fig. 3.

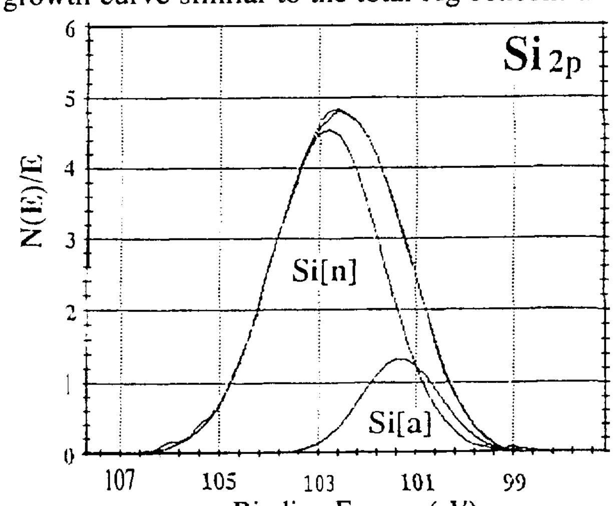


Figure 5. A typical Si 2p XPS spectrum with two components, Si[n] and Si[a], at sample temperature 300 ‘C.

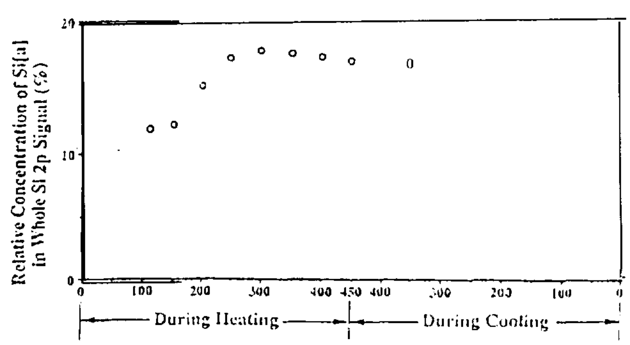


Figure 6. The relative concentration of Si[a] in the whole Si 2p signal at various temperatures.

The corresponding changes in oxygen are clearly observed after the deconvolution of asymmetric O 1s signals. Fig.7 shows two typical deconvoluted O 1s XPS spectra, (a) recorded at 50 CC, (b) obtained at 300 °C. It was observed that at low temperatures (below 100 °C during heating and below 220 °C during cooling) the XPS spectra of O 1s fitted to the sum of two

components. According to the binding energy of the peak, the peak with high binding energy was assigned to the bridging oxygen (BO, Si-O-Si linkage) and the peak with low binding energy to the non-bridging oxygen (NB O, Si-O- M, where the M refers to metal including Ag or an oxygen vacancy where oxygen does not bond to Si). However, with the temperature increase from 110 ‘C to 450 “C, a third peak with lower binding energy must be added to the XPS spectra of O 1s in order to make a best fit. This trend continues during cooling until the temperature is lower than 220 “C. This peak was also attributed to the non-bridging oxygen and is designated NBO* hereafter. The relative concentrations of NBO and NBO* against sample temperature are plotted in Fig. 8. It is clearly seen that the NBO increases until 350 “C and then drops continuously during heating and cooling until 50 ‘C. The NBO concentration returns to the original value at temperatures below 50 “C. The NB O* appeared at sample temperatures above 110 ‘C during heating, gradually decreased until 400 ‘C, and finally disappeared at 220 °C during cooling.

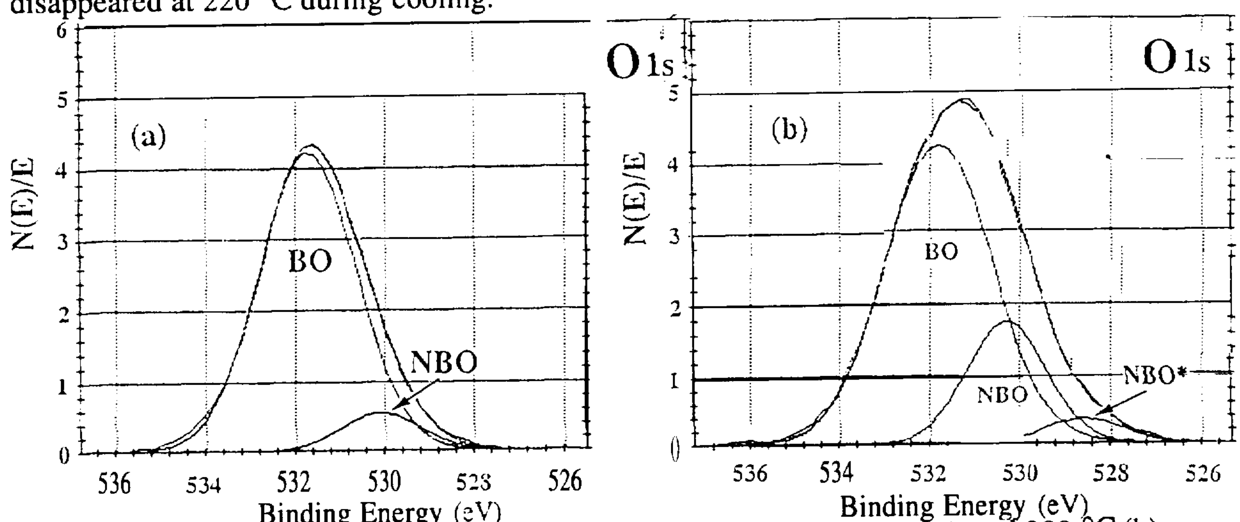


Figure 7, The deconvoluted O 1s spectra at sample temperature 50 ‘C (a) and 300 °C (b).

IV. DISCUSSION

A. Silver diffusion during annealing

Silver from the ion-exchanged sample diffuses toward the surface and forms neutral silver during heat treatment as clearly seen in Figs, 2 and 3. Diffusion toward the surface and precipitation of silver are caused by thermal relaxation of the surface tensile stress introduced by the size difference between Ag^+ and Na^+ [25] (the ratio of ionic radii, r_{Ag^+}/r_{Na^+}, *is* 1.29) [26] during the ion exchange process. During the ion exchange, larger Ag^+ ions (depending on their concentration excess and the decrease of the glass viscosity due to the high temperature) enter the glass structure and replace smaller Na^+ ions. After cooling some Ag^+ ions are consolidated into the glass structure, but tensile stress introduced by the size difference between Ag^+ and Na^+ is still there.

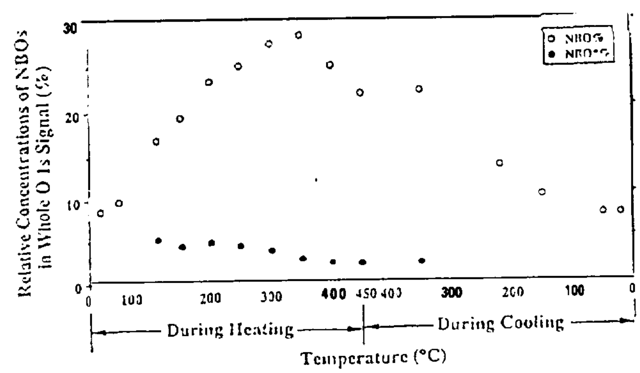


Figure 8. The relative concentration of NBO and NBO* in the whole O1s signal at various sample temperatures.

At low temperature during heating, only a few Ag^+ ions acquire enough energy to overcome the static barrier potential (i. e., bonding energy of Ag-O) produced by oxygens bonded to them and move towards the still considerably stressed surface. Once the temperature is high enough (above 100 °C), there are more activated Ag^+ ions moving toward a more relaxed surface resulting from precipitation or colloid formation of neutral silver as discussed in section C. So in this stage, the total silver concentration on the glass surface increases at a very high rate. As more silver segregates at the surface, a repulsive potential is generated which

retards the further diffusion of silver. Hence the enrichment of silver on the surface apparently slows down at temperatures above 350 'C. Early studies of stress release in soda-lime glasses showed that a faster release rate occurred at sample temperatures between 150 and 250 °C [27,28]. Our Ar^+ sputtering profile study shows that in ion-exchanged glass samples silver reaches a depth of -2 μm. Hence, stress release in the sample similar to that in soda-lime glass is expected. More recently, Kern et al [29] and Shintani *et al* **[30]** have found that the tensile stress of SiO_2 film decreases with the increasing phosphorous concentration. Therefore, more Ag^+ ions diffuse to surface and precipitate there under heating may also reduce the tensile stress inside the ion-exchanged layer. Consequently, a surface diffusion of silver between 100 °C and 350 'C is a reasonable result.

B. Silver chemical state change during heating and cooling

The formation of silver colloids in the ion-exchanged sample during heating also result from the differences in Gibbs free energies of pure silver, silver oxide and SiO_2. Even though the Gibbs free energy of silver oxide (-2.68 Kcal/mole at 25°C) is lower than that of pure silver (O Kcal/mole at 25°C) and higher than that of SiO_2 (~ -200 Kcal/mole at 25°C), the dissociation of Ag-O bonds to form Si-O and Ag-Ag bonds stall results in a net loss in the system energy [31]. Combining this energy loss with the reduction of tensile stress, the thermodynamics of the surface diffusion of silver is clear. Thus, the silver inside the glass network will move to the surface and precipitate during heat treatment in order to maintain the system in a minimum energy state. Similar silver colloid formation in silver-implanted silica after annealing was previously observed and studied by infrared spectroscopy and transmission electron microscopy [7,32].

Combining the changes of neutral and oxidized Ag, as shown in Fig. 3, and the corresponding changes for Si and O, as shown in Figs. 6 and 8, with the phenomenon of neutral silver precipitates during heating, a disappropriation reaction mechanism for Ag^+ on the surface can be deduced. That is, for Ag^+ ions moved to the surface the following reaction may have taken place at the high temperature:

$$2Ag^+ \longleftrightarrow Ag + Ag^{2+} \quad (1)$$

According to the literature [33], the binding energies of the silver species have the following order: $Ag^{2+} < Ag^+ < Ag$. So if the reaction (1) does take place, the relative concentration of Ag^+ will decrease and the relative concentration of Ag" will increase. However, it is clearly seen in Fig. **3** that the concentration of Ag^+ does not decrease at temperatures between 110 and 300 'C during heating. And it does decrease from 350 to 450 "C during heating. It should be noted that the concentration changes of Ag^+ during heating result from the variations in both Ag^+ and Ag^{2+} species because Ag^{2+} cannot be distinguished in the Ag spectra. Therefore, even though the forward disappropriation reaction occurred, the small concentration increase in the Ag^+ at temperatures between 110 and 300 CC during heating results from the fact that the production rate of Ag^{2+} is less than the accumulation rate of Ag^+ on the surface. It is obvious that the reverse disappropriation reaction occurs during cooling since the concentration of the Ag^+ increases during the cooling process.

Moreover, the appearance of NBO* and Si[a] also supports this deduction. Apparently, there are more negative charges on the oxygens around Ag^{2+} ions than on the oxygens around Ag^+ ions. Consequently, the effective charges on Si and NBOs will decrease, which causes the formation of Si[a] and NBO*, $\equiv Si\text{-}O^{2-} \cdots Ag^{2+}$. It is well-known that the disappropriation reaction is conditionally sensitive. In the present case, the condition is the temperature. As the reaction temperature goes down, the reaction will move in the opposite direction. Hence, during the cooling, the relative concentration of Ag^+ increases and the neutral Ag content decreases, With the reverse disappropriation reaction, the Ag^+ ions increase and the Ag^{2+} ions decrease; the NBO* and Si[a] will disappear as shown in Fig. 6.

C. Responses of silicon and oxygen during annealing

The silicon surface concentration does not change during annealing as shown in Fig. 4 but two species of Si were observed during annealing. The major Si species with high BE is the

that silver colloids are formed through the relaxation of the tensile stress and the reduction in Gibbs free energies of pure silver, silver oxide and SiO_2.

ACKNOWLEDGMENTS

This work was supported in part by the Materials Research Center of Excellence at University of Texas at El Paso under NSF contract number HRD-9353547.

VI. REFERENCES

[1] For example, N.F. Borrelli and D.W. Hall, "Nonlinear Optical Properties of Glasses", pp. 87-124 in *Optical Properties of Glass*, edited by D.R. Uhlmann and N.J. Kreidl, Am. Ceram. Sot., Westerville, OH, 1991.

[2] S. Schmitt-Rink, D.A.B. Miller, and D.S. Chemla, *Phys. Rev.*, B35 ,8113 (1986) ,

[3] C.F. Flytzanis, F. Hache, M.C. Klein, D. Ricard and P. Roussignol, *"Nonlinear Optics in Composite Materials"*, in: Progress in Optics, ed. E. Wolf, vol. 29 (Elsevier, New York, 1991), p.321.

[4]G.W. Arnold and J.A. Borders, *J. Appl. Phys.* 48,1488 (1977).

[5]N.D. Skelland, P.D. Townsend, *J, Non-Cryst. Solids, 188,243 (1995),*

[6] R.H. Magruder III, R.A. Weeks, S.H. Morgan, Z. Pan, D.O. Henderson, R.A. Zuhr, *J. Non-Crys. Solid,* 192&193, 546 (1995).

[7] S. F/. Houde-Walter and B.L. McIntyre, *J, Non-Crys. Solid, 107,316 (1989).*

[8] *G.* De, Marchi, F. Gonella, P. Mazzoldi, G. Battaglin, E.J. Knystyautas, and C. Meneghini, *J. Non-Cryst. Solids, 196,79 (1996).*

[9] G. De, A. Licciulli, C. Massaro, L. Tapfer, M. Catalano. G. Battaglin, C. Meneghini, and P. Mazzoldi, *J. Non-Crys. Solids,* 194,225 (1 996).

[10] H. Hofmeister, M. Dubiel, H. Go, and S. Thiel, *J. Microscopy,* 177,331 (1995).

[11] L. Maya, M. Paranthaman, T. Thundat, and M.L. Bauer, *J. Vat. Sci.Technol.*,B14, 15 (1996).

[12] S.N. Houde-Walter and B.L. McIntyre, *J. Non-Crys. Solid*, 107, 316 (1989),

[13] K. Fukumi, K. Kageyama, and M. Satou, *J. Mat. Res.*, 10,2418 (1995),

[14] N.A. Sharaf, R.A. Condrate, and A.A. Ahmed, *Mat. Lett.*, 11,115 (1991).

[15] C.F. Flytzanis, F, Hache, M.C. Klein, D. Ricard and P, Roussignol, *"Nonlinear Optics in Composite Materials"*, in: Progress in Optics, ed. E. Wolf, vol 29 (Elsevier, New York, 1991), p.321.

[16] D.E. Day, "Mixed Alkali Glasses-Their Properties and Uses", *J. Non-Cryst. Solids, 21,343 (1976).*

[17] A.J. Burggraaf and J. Cornelissen, Phys. *Chem. Glasses* ,5, 123 (1964).

[18] P.W. Wang, *J. of Vac. Sci. and Tech.*, A14, *465 (1996).*

[19] S.N. Houde-Walter, J.M. Inn-m, A.J. Dent, and G. F/. Greaves, *J. Phys. Chem.*, 97,9330 (1993).

[20] R. Bruckner, H.-U. Chun and H. Goretzki, *Glastechn. Ber.*, 51, 1 (1978).

[21] J.S. Jen and M.R. Kalinkowski, *J. Non-Cryst. Solids, 38&39, 21 (1980) .*

[22] P.W. Wang, L.P. Zhang, L. Lu, D.V. LeMone and D.L. Kinser, *Appl.* Surf. *Sci.*, 84,75 (1995).

[23] G.M. Renlund, S. Prochazka, and R.H. Doremus, *J. Mater. Res.*, 6,2723 (1991) .

[24] P.W. Wang, L.P. Zhang, L.Lu, D.V. LeMone. and D.L. Kinser, *Appl. Surf. Sci., 8,75* (1995) .

[25] P.W. Wang and L.P. Zhang, *J. Non-crystal. Solids, 194, 129 (1996).*

[26] "CRC Handbook of Chemistry and Physics", edited by R. C. Weast and M.J. Astle, (CRC Press Inc., Boca Raton, 1981), p. F-216.

[27] A. F. Van Zee and H. M. Noritake, *J. of Am. Ceram. Sot., 41,164 (1958).*

[28] S.S. Kistler, *J. of Am. Ceram. Society, 45,59 (1962).*

[29] W. Kern, G.L. Schnable, and A.W. Fisher, *RCA Rev., 37 (1976).*

[30] A. Shintani, S. Sugaki, and H. Nakashima, *J. Appl. Phys., 51,4197 (1980).*

[31] "CRC Handbook of Chemistry and Physics", edited by D.R. Lide, (CRC Press Inc., Boca Raton, 1991) p. 5-38 and 5-49.

[32] Yiping Feng, Paul W. Wang, and Walter L. Roth, Mar. *Res.* Sot. *Proc., 135, 535 (1989).*

[33] J.S. Hammond, S.W. Gaarentroom, and N. Winograd, *Annal. Chem..*, 47, 2193 (1975).

Page intentionally left blank

/V ASA

52 / 51

URC97133

The Integration of Geographical Information Systems and Remotely Sensed Data to Track and Predict the Migration Path of the AfricanizedHoney Bee

Charles Ward, Jessica Bravo, Rosalia De Luna,
Gerardo Lopez, Itza Pichardo, Danny Trejo, and Gabriel Vargas

Pan American Center for Earth and Environmental Studies (PACES)
University of Texas at El Paso
El Paso, Texas 79968

Introduction

One of the research groups at the Pan American Center for Earth and Environmental Studies (PACES) is researching the northward migration path of AfricanizedHoney Bees or often referred to in the popular press as killer bees. The goal of the Killer Bee Research Group (KBRG) is to set up a database in the form of a geographical information system, which will be used to track and predict the bees future migration path. Included in this paper is background information on geographical information systems, the SPANS Explorer software package which was used to implement the database, and Advanced Very High Resolution Radiometer data and how each of these is being incorporated in the research. With an accurate means of predicting future migration paths, the negative effects of the Africanizedhoney bees maybe reduced.

History of Africanized Honey Bees

In 1956, the African honey bee was introduced into Brazil. The hope was to create a strain of honey bee which was better suited for the tropical environment of South America and thereby improving honey production. In 1957, twenty six bee colonies escaped and began interbreeding with the native bees. These new AfricanizedHoney Bees (AHB) have a greater tendency to swarm and abscond than common European honey bees. Since their initial release in Brazil, the AHB have been spreading at a rate of 80 to 500 kilometers annually. In the path of their northward migration, they have reduced honey production and negatively impacted crop pollination, livestock production, tourism and public health. The migration of AHB reached the United States in October of 1990 near Hidalgo, Texas. Currently, AHB have been sighted in south central Texas, southern New Mexico and Arizona and the central valley of California. A more thorough study of AHB is reported in [4].

Geographical Information System

A geographical information system (GIS) can be defined as "a computer-based information system that enables capture, modeling, manipulation, retrieval, analysis and presentation of geographically referenced data" [7]. One of the best features of using a GIS is that it provides the user a way to effectively visualize spatial information. A GIS gives a visual reference point, which helps the user to understand.

GIS use two basic types of data : geographic and attribute. Geographic data is composed of locational or spatial information about each specific feature or entity of data. Attribute data provides non-spatial information describing the characteristics or properties of each entity [1].

There have been many similar applications of GIS. An example is the Center for Health Applications of Aerospace Related Technologies (CHAART) at NASA Ames Research Center, in which the risk of Lyme disease transmission was studied [2].

The GIS software currently being used by the PACES KBRG is SPANS Explorer, which is manufactured by TYDAC Technologies Inc. SPANS Explorer comes in several software formats. The PACES KBRG is running a Windows 95 version of the program on Pentium 100 MHz computers.

Data Layers

The foundation for the data layers is a base map. Information can then be added to the base map by overlaying data layers, which can be displayed using different graphical formats. These formats are point, line, area, raster and quadtree. Point data represents locations of geographic phenomena, such as elevation or a sampling site. Line data consists of a set of connected points. Some examples of line data include rivers, roads and non-physical boundaries, such as county and state lines. Area data consists of a region enclosed within a boundary. Our research group is storing our information in this format. The data we have collected, so far, is in the format of number of AHB sightings in a county and hence is entered as area data. Raster data is a grid data structure which represents an area of the Earth's surface as a matrix of equal sized cells. Aerial photography, satellite and radar imagery are examples of raster data. Quadtree data is a more sophisticated form of raster data. Quadtree data represents an area of the Earth's surface as a matrix of variable sized cells. Areas with more detail are composed of smaller cells to preserve the resolution. For example, a quadtree of a coastal area uses large cells to represent the ocean and smaller cells to represent the greater detail of the coastline.

Methods of Entering Data

Data can be entered into SPANS Explorer manually or through the importation of other file formats. Attribute data can be added to a data layer in a manner similar to using a common spreadsheet. This is the method that the research group has used. Geographical data, such as boundary lines, can be input with the use of a mouse. SPANS Explorer also allows for the importation of a large variety of data formats. It comes with a raster and vector translator companion programs which can translate formats such as United States Geological Survey (USGS) 7.5 Minute Digital Elevation Model.

User Interfaces

SPANS Explorer has querying and charting functions that allow for the analysis of information. The software package has two methods for performing queries of the data layers, spatially and non-spatially. A spatial query involves the analysis of data in only a user defined area of the base map. The program allows a specific area to be chosen for query, by enclosing it in a circle, rectangle or polygon. A non-spatial query involves an analysis of the data layer as a whole. A query is performed by defining an equation, An equation is made up of one or more statements, which consists of an attribute and a constant value separated by a relational operator. Several statements can be combined with the use of a logical operator, such as AND or OR. When a query is executed the entities, which fit the conditions of the defined equation, are highlighted. The query function has been used by the research group to observe which counties in Texas have had a continuous increase in the AHB population. Data contained in a data layer can also be analyzed in the form of a chart. A chart can be created by highlighting data in an open data layer. The data can be seen in a line, bar, pie or radar chart format.

Attribute Data Used in AHB Migration GIS

The present attribute data that is being used in the AHB migration GIS can be grouped into the categories of weather, land usage and sightings. The data in each of these categories are being added to the GIS in the form of data layers.

The first category, weather, consists of data on the number of frost free days, the average monthly temperature and amount of precipitation in an area(county). All three have been shown to influence the rate and possible migration paths of AHB. This information is being acquired from the National Climate Data Center (NCDC).

Another type of data that influences AHB migration is land resource usage. This indicates if land is a wilderness area or if the land is currently being used for agriculture. A source of this data is still being searched for at the present. An alternative approach to filling this layer is to use the results of a classification model driven by remotely sensed data.

The final information that is being used in the set of data layers is the sightings of AHB swarms by year. AHB sightings give the important information of where the next migration move will start from, The best source of this information, to this point, has been "Detection Records of AfricanizedHoney Bees in Texas During 1990, 1991 and 1992" [3]. This paper gives the estimated number of trapped and free swarms of AHB in Texas, over those three years.

Remotely Sensed Data

Another form of data that the research group will be integrating into the AHB migration GIS is remotely sensed data from Advanced Very High Resolution Radiometers (AVHHR). AVHRR is a broad-band, four or five channel (depending on the model) scanner that senses in the visible, near infrared, and thermal infrared portions of the electromagnetic spectrum. The spectral ranges are as follows :

Channel #	Wavelength in micrometers
I	0.58-0.68
2	0.725-1.10
3	3.55-3.93
4	10.3 -11.3
5	11.5 -12.5

The AVHRR scanner is carried on the National Oceanic and Atmospheric Administration (NOAA) Polar Orbiting Environmental Satellites (POES). The POES orbit the Earth 14 times each day at an altitude of 833km(517mi). The data has a ground resolution of approximately 1.1 km. Four out of every five samples along the scan line are used to compute one average value and the data from only every third scan line are processed. This yields a 1. 1km by 4-km resolution at nadir. One problem experienced by the AVHRR scanner is clouds. Several AVHRR overpasses are required to ensure cloud-free images. The duration of consecutive daily observations is called the compositing period. An image which provides a clear observation of a large ground surface area at reasonable nadir viewing angles is included in the composite.

AVHRR data is used to calculate the Normalized Difference Vegetation Index (NDVI) which gives the greenness or biomass of an area. The NDVI is calculated by taking the difference between the near-infrared (AVHRR Channel 2) and visible (AVHRR channel 1) reflectance values and dividing by the total reflectance. This calculation results in a value between -1 and 1, where a negative value tends to represent clouds, snow, water and other non-vegetated surfaces; positive values represent vegetated surfaces. The value is then scaled to fall between O and 200, and stored as an 8 bit number. When compositing the image, the highest value for a pixel during that period is used.

NDVI data has a problem distinguishing between clouds and water because the value over water is much lower than that of cloud. As a result of this cloudy images are taken over bodies of water. To solve this problem values less than 100 are flipped so clear observations will be chosen. For example a value of 99 becomes a value of 1.

NDVI data has a problem distinguishing between clouds and water because the value over water is much lower than that of cloud. As a result of this cloudy images are taken over bodies of water. To solve this problem values less than 100 are flipped so clear observations will be chosen. For example a value of 99 becomes a value of 1.

AVHRR data has been collected is performed by the EROS Data Center by the USGS, since May 1987. The KBRG has acquired the AVHRR data for the Conterminous United States from the USGS for the year 1989 to 1995.

From various statistical analyses, it is believed that a correlation maybe found between the NDVI data and the AHB migration paths. Preliminary research has found that there tends to be a difference between the standard deviation of the NDVI values for a given county that has Africanizedbees and those that do not. The NDVI data set contains statistical information, including the mean, median, maximum, minimum and standard deviation for every county in the conterminous United States. A classification scheme is being developed to determine whether a county is likely to be infested with bees. The correlation coefficient is being calculated between a group of test counties. The county is then classified based on whether it tends to be more closely related to counties which are known to have AHB or those that do not. If ibis can be done, a profile based on NDVI data can be created to predict future AHB migration. Results of the work to date are provided in [5].

Summary and Future Work

The KBRG has use a GIS to help track, analyze and predict the migration path of the AHB. The group continues to collect information to improve the quality of the predictions. Remotely sensed data, specifically AVHRR data, is being integrated into the GIS and analyzed by other programs to determine any correlation can be found with the AHB migration path. Correlation studies based on other measures associated with NDVI will be investigated to expand the predictive models.

Acknowledgments

We wish to thank NASA for its support of this project through the grant NCCW-0089. Also, we wish to extend gratitude to Byron Wood and Louisa Beck of the Center for Health Applications of Aerospace and Related Technologies at NASA Ames Research Center for their assistance in the formulation of this project.

References

[1] Burke, Diane, SPANS Explorer: The Future of Geographic Analysis Defined, TYDAC Research Inc. ,1996.

[2] Dister, S. W., L.R. Beck, B. L.Wood, R. Falco, and D. Fish, "The use of GIS and remote sensing technologies in a landscape approach to the study of Lyme disease transmission risk," Proceeding. GIS '93: Geographic Information Systems in Forestry. Environmental and Natural Resource Management, Vancouver, B. C., Canada, pp. 15-18 February 1993.

[4] Navarro, Hector, "An Investigation of the Migration of Africanized Honey Bees into the Southern United States", Proc. of the 1997 NASA URC-TC '97 National Technical Conference on Education, Aeronautics, Space, Autonomy, Earth, and Environment, Albuquerque, NM, Feb. 1997.

[5] Ward, Charles, "A System To Characterize Landscape Phenology Using AVHRR Data", MSEE Thesis, The University of Texas at El Paso, December 1996.

[6] Wood, Byron, Louisa Beck, Sheri Dister, Brad Lobitz, and Vince Ambrosia, Readings in Remote Sensing, Geographic Information Systems, and Related Technologies, Center for Health Applications of Aerospace Related Technologies, NASA Ames Research Center, 1995.

[7] Worboys, Michael F., GIS : A Computing Perspective, New York : Taylor& Francis Inc., 1995.

Page intentionally left blank

URC97134 // ... 2133/27

SURFACE STATES AND EFFECTIVE SURFACE AREA ON PHOTOLUMINESCENT P-TYPE POROUS SILICON

S.Z. WEISZ, A. RAMIREZ PORRAS, AND O RESTO,

Department of Physics, University of Puerto Rico, Rio Piedras, PR 00931

and

Y. GOLDSTEIN, A. MANY, AND E. SAVIR

Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel

ABSTRACT

The present study is motivated by the possibility of utilizing porous silicon for spectral sensors. Pulse measurements on the porous-Si/electrolyte system are employed to determine the surface effective area and the surface-state density at various stages of the anodization process used to produce the porous material. Such measurements were combined with studies of the photoluminescence spectra. These spectra were found to shift progressively to the blue as a function of anodization time. The luminescence *intensity* increases initially with anodization time, reaches a maximum and then decreases with further anodization. The surface state density, on the other hand, increases with anodization time from an initial value of ~2x1 0^{12} cm^{-2} for the virgin surface to -1013 cm^{-2} for the anodized surface. This value is attained already after -2 min anodization and upon further anodization remains fairly constant. In parallel, the effective surface area increases by a factor of 10-30. This behavior is markedly different from the one observed previously for n-type porous Si.

INTRODUCTION

The work presented here was motivated by the possibility of the use of porous silicon as a spectral sensor. Porous silicon, [1-4] (PS) obtained by electrochemical etching procedures applied to crystalline Si surfaces, when illuminated by u.v. light, exhibits high luminescence efficiencies in the visible range. This effect, and the parallel electroluminescence effect, promise the possibility of the realization of PS based optoelectronic devices on top of crystalline silicon. One such possibility is the use of porous silicon on top of silicon charge coupled devices (CCD). Silicon CCD devices are used in many applications of optical imaging. However, the silicon spectral sensitivity is quite limited in the blue and u.v. because of surface recombination. The incorporation of porous silicon into a silicon based imaging system may enable the extension of the spectral range towards the u.v.

The present work concentrates on gaining further insight into the porous silicon photoluminescence process and the role of the surface in this process. To that end we have employed combined studies of the luminescence spectrum, the surface-state density and the effective surface area of the porous surface. Such studies were carried out at different stages of the anodization process and thus for different morphologies of the porous surface. The luminescence spectra were measured by conventional methods. The surface state characteristics and the effective surface area were determined by pulse measurements on the PS/electrolyte system. This system is particularly suitable since a capacitative contact to the terrain of the porous surface is best achieved by an electrolyte, and it was successfully used[6] to investigate n-type PS. There we found[6] a strong correlation between the surface-state density near the conduction band edge and the luminescence intensity. In this paper we present similar measurements on *p-type* porous Si and we compare the results with those obtained on n-type material.

EXPERIMENTAL

The starting material was high-grade p-type silicon wafers of resistivity in the range 20–50 Ωcm. A p^+ layer was formed by diffusing metallic Al into one of the faces to obtain an ohmic contact. The sample was attached to a cylindrical Teflon cell via a Kalrez O-ring, the sample constituting the bottom of the cell, with its free surface facing upwards. The samples were etched in 20°/0 HF. In order to prepare the porous surface,[4] a solution of HF, ethanol and water (1:1 :2) was poured into the cell. A platinum electrode was immersed in the solution and a spring contact was attached to the p^+ contact. The anodization of the Si surface was carried out with a current density of 100 mA/cm^2.

The luminescence of the PS was excited by a 10 mW He-Cd laser beam (λ = 442 rim). The luminescence spectra at different stages of the anodization process were measured by a Control Development spectrometer.

The electronic characteristics of the PS/electrolyte interface were studied at different stages of the anodization process, starting from the "virgin" surface and up to an anodization time of 20 minutes. To that end, the anodizing solution was replaced after each anodization stage by an electrolyte; an aqueous solution of KCl. The measurement technique applied to the semiconductor/electrolyte (WE) interface has been described elsewhere,[5] and will be reviewed only briefly here. A short voltage pulse of duration T = 20 µsec, applied between the Pt electrode and the sample's p^+ contact, is used to charge up the interface region. The voltage drop across these electrodes, measured just after the termination of the pulse (T + dT), represents to a very good approximation the change δV_s in barrier height across the semiconductor space-charge layer induced by the applied pulse. If an insulating layer, such as an oxide, is present at the semiconductor surface, the measured voltage drop is $\delta V_s + \delta V_g$, where δV_g is the drop across the insulating layer. Obviously, $\delta V_g = Q_{tot}/C_g$, where Q_{tot} is the total charge density induced at the surface and C_g is the "geometric" capacitance of the insulating layer (per cm^2). Q_{tot} is obtained from the voltage V_c developed across a large series capacitor C, again at the termination of the pulse. Pulses of varying amplitude are applied singly, one per data point taken. In this manner possible damage to the porous surface is minimized,

In general, Q_{tot} is made up of three components:

$$Q_{tot} = \delta Q_{sc} + \delta Q_{ss} + Q_L \ , \tag{1}$$

where δQ_{sc} is the change in the free space-charge density, δQ_{ss} is the change in surface-state charge density, and Q_L is the charge density that has leaked across the interface due to imperfect blocking of the S/E interface. In order to determine each component of Q_{tot}, the platinum electrode is shorted to ground by an electronic switch at T + dT, where dT is very short (O. 1 – 0.2 µsec), just sufficient to permit accurate readings of δV_s and V_c right after the termination of the pulse. At this point, charge redistribution between C and the S/E interface begins to take place. In the first stage, the free charge δQ_{sc} and its equal counterpart in C discharge relatively fast through the low resistance of the sample and the electrolyte. The decay constant associated with this process is typically several microseconds. As a result, V_c decays to the value $\delta Q_{ss}/C$, δQ_{ss} being the charge remaining in C after the fast decay process. Thereafter, V_c decays to zero usually much more slowly, as charge trapped in the surface states by the charging pulse are thermally re-emitted into the conduction band, in the case of n-type semiconductor, or valence band, in the case of p-type semiconductor. The decay time is larger the farther away the surface states are located energetically from the relevant band edge and the lower the temperature.[7] If charge leakage exists, V_c does not decay to zero but to the value Q_L/C. Subsequently it remains practically constant since the leaked charge has been lost from the interface and the remaining charge Q_L on C cannot be dissipated. This behavior enables the separate determination of the three components in Eq. (1), all as functions of δV_s. In what follows, we shall express these components in terms of hole surface densities: $\delta P_s = \delta Q_{sc}/q$ and $\delta P_{ss} = \delta Q_{ss}/q$, where q is the absolute magnitude of the electronic charge.

In most cases a space-charge layer already exists at the semiconductor surface, before applying any bias. It is characterized by an equilibrium barrier height V_{so} and an equilibrium surface hole density P_{so}. If surface states are present, there may also be an equilibrium density P_{sso} of "occupied" surface states. V_{so} and P_{so} can be determined quite accurately from measurements in the depletion range.[8] The entire plots of P_s and P_{ss} vs. V_s can then be constructed by using the relations $V_s = V_{so} + \delta V_s$, $P_s = P_{so} + \delta P_s$, and $P_{ss} = P_{sso} + \delta P_{ss}$. So much so in the absence of an insulating layer (such as an oxide) at the semiconductor surface. If such a layer is present, the as-measured barrier height, i. e., the measured voltage drop between the Pt electrode and the p^+ contact just after the pulse termination, yields V_s+V_g, where V_g is the voltage drop across the insulating layer.

RESULTS AND DISCUSSION

In Fig. 1 we present typical photoluminescence spectra of p-type PS surfaces prepared by anodization. The different anodization times are marked on the spectra. We notice that the luminescence intensity at the beginning increases with anodization time, attains a maximum and then decreases. This behavior is illustrated by the higher curve in Fig. 2 and is quite similar to that observed[6] previously for n-type PS. The lower curve in Fig. 2 shows the appreciable blue shift of the spectra, suggesting that, on the average, the porous structure gets finer with anodization time.

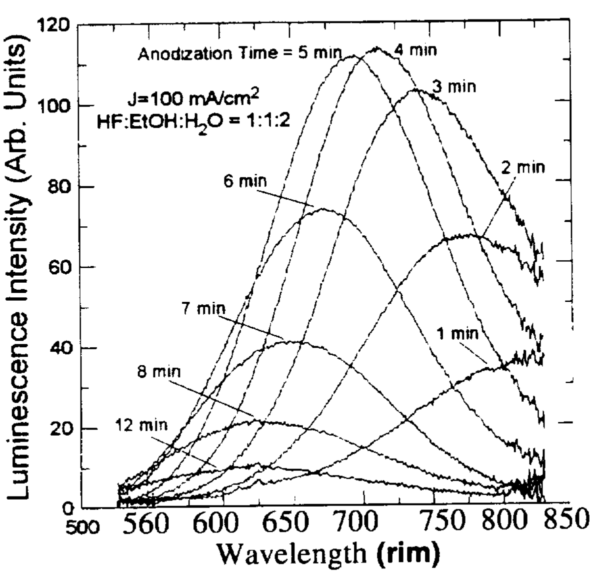


Fig.1. Photoluminescence spectra for various anodization times, as marked.

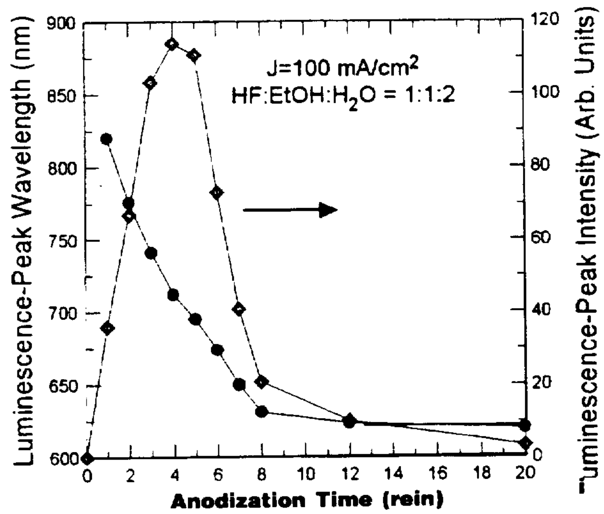


Fig. 2. Peak-photoluminescence intensity and wavelength as functions of anodization time for the sample in Fig. 1.

Typical results of the free surface-hole density P_s (diamonds) and the density of occupied surface states P_{ss} (stars) against the as-measured barrier height V_s+V_g as obtained for an etched virgin silicon surface, are displayed in the semilog plot of Fig. 3. In the depletion range, P_s is negative but, because of the logarithmic scale used, the plot is that of $-P_s$. The light curve, in the accumulation range, labeled $C_g = \infty$, represents the theoretical dependence of P_s on V_s for an oxide-free surface ($C_g = \infty$, $V_g = O$), as derived from a solution of Poisson's equation for the value of the hole bulk concentration p_b marked in the figure. It is seen that this curve does not account well for the data in the accumulation range. The best fit, represented by the bold curve, labeled $C_g = 4.5\ \mu F/cm^2$, was obtained by assuming the presence of an oxide layer of $C_g = 4.5\ \mu F/cm^2$, corresponding to an oxide thickness of -4 A, Again, this behavior is similar to that found on n-type PS. Turning now to the surface-state hole occupancy P_{ss}, it is seen to rise slowly from a low

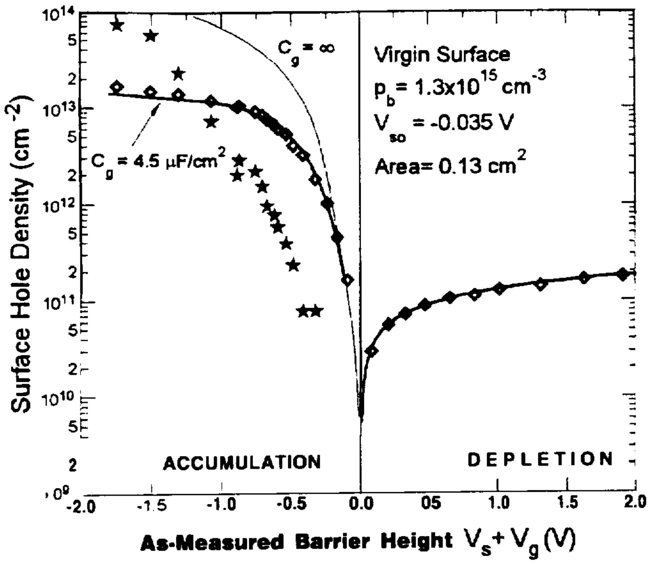


Fig. 3. Free surface hole density P_s (diamonds) and density of occupied surface states P_{ss} (stars) vs. the as-measured barrier height V_s+V_g for a virgin Si surface. The light and bold curves are theoretical plots of P_s as explained in the text.

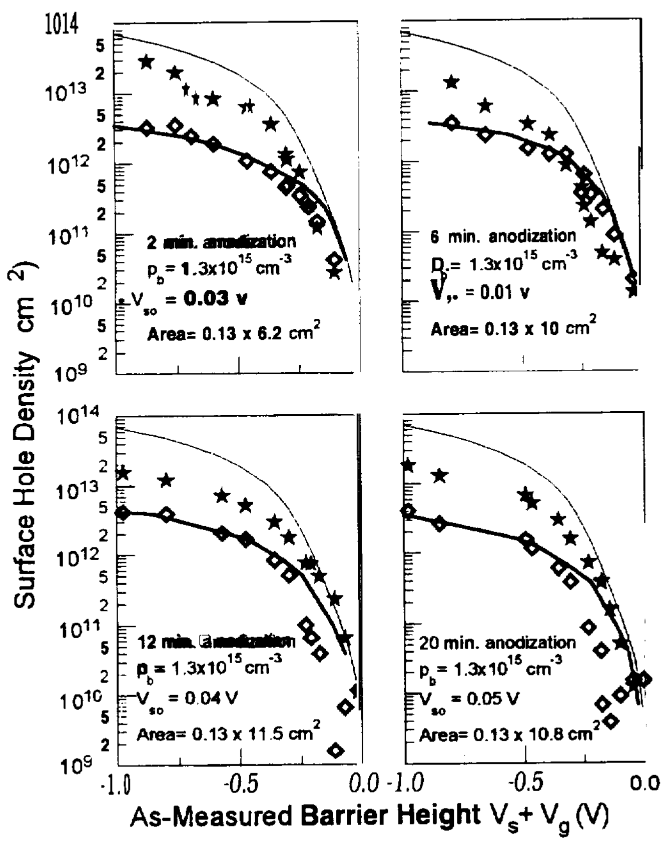


Fig. 4. P_s (diamonds) and P_{ss} (stars) vs. V_s+V_g for four anodized, porous Si surfaces. The light and bold curves are theoretical plots of P_s.

value of -10^{11} cm^{-2} at $V_s+V_g \approx -0.5$ V up to $\sim 10^{14}$ cm^{-2} at $V_s+V_g \approx -2$ V. No saturation value for the surface-state density could be reached. This is quite different than the behavior we found[6] on n-type PS, where the surface-state density saturated at 10^{12} cm^{-2}. Because of this lack of saturation, we shall use the values of P_{ss} at $V_s+V_g \approx -1$ V to compare surface state occupancies for different anodization times. For the case of the virgin surface, this value is ~2x1 012 cm^{-2}.

Results of P_s and P_{ss}, similar to those in Fig. 3, for four porous surfaces are presented in Fig. 4. These results were obtained after the sample of Fig, 3 has been anodized for different times, as marked in the figure. Since our aim is to compare the surface-state densities, we show only the accumulation range here. Because of the increase of the effective surface area, the highest surface potential barriers (for holes) attained were around -1 volt. The curves in the figure are theoretical, calculated for the same C_g values as in Fig. 3. We notice again that for all four porous surfaces, the surface-state density (stars) increases monotonously with the potential barrier through the whole region shown and does not exhibit signs of saturation. As mentioned above, we choose for comparison the values of P_{ss} at --1 volt. These values in the figure are scattered around 2x1 0^{13} cm^{-2}, about an order of magnitude higher than on the virgin surface. However, the interesting thing is that this surface-state density remains fairly constant with anodization time, very much different from the results found[6] for n-type PS. A behavior similar to that of the surface states was observed also for the effective surface area. In Fig. 5 the surface area factor, e.g. the ratio of the effective area to the area of the virgin sample, for two typical samples is plotted against the anodization time. The area factors were derived from measurements of the type shown in Figs. 3 and 4. The results for the two samples are qualitatively the same; the area factor increases with anodization time till it reaches a saturation. The saturation values observed varied from sample to sample

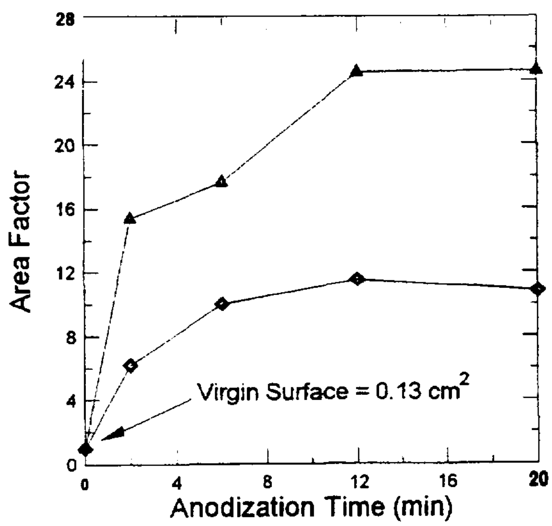


Fig.5. Surface area factor vs. anodization time for two samples.

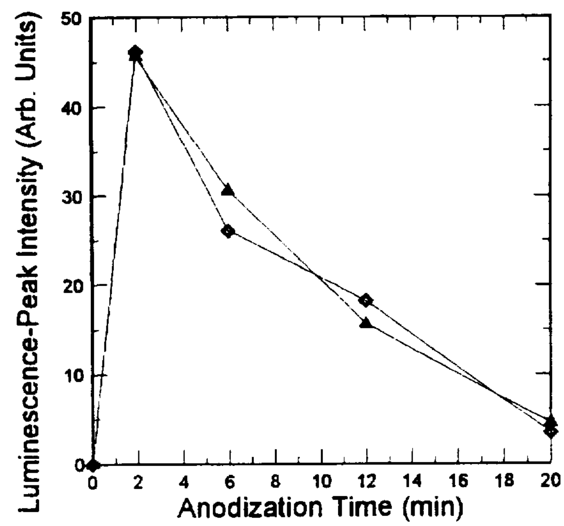


Fig. 6. Photoluminescence-peak intensity vs. anodization time for two samples.

between 10 and 30.

In parallel with the surface-state density, we measured the photoluminescence spectrum for each anodization time. These measurements were performed after the measurements on the PS/electrolyte system. In Fig. 6 the luminescence- peak intensity is plotted against the anodization time for the same two samples as in Fig. 5. We notice that the maximum luminescence peak obtains already after 2 min anodization and then the luminescence decreases upon further anodization. In Fig. 7 we plot the value of the luminescence-peak wavelength as a function of anodization time. The peak shifts to the blue upon anodization close to 200 nm from its value at 2 min anodization.

Comparing the results of Figs. 6 and 7 to those in Fig. 2 we notice that they both show comparable blue shifts, but there is a discrepancy in the anodization time needed to attain maximum luminescence. To check the influence of the KCl electrolyte used in the electronic measurements, we measured the spectra on PS samples that after anodization were immersed for 10 min into a KCl electrolyte. The results are shown in Fig. 8 and, for comparison, we also re-

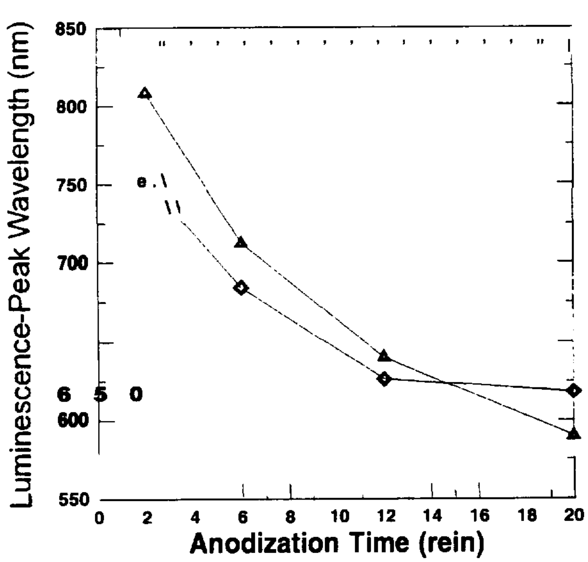


Fig. 7. Photoluminescence-peak wavelength vs. anodization time for two samples.

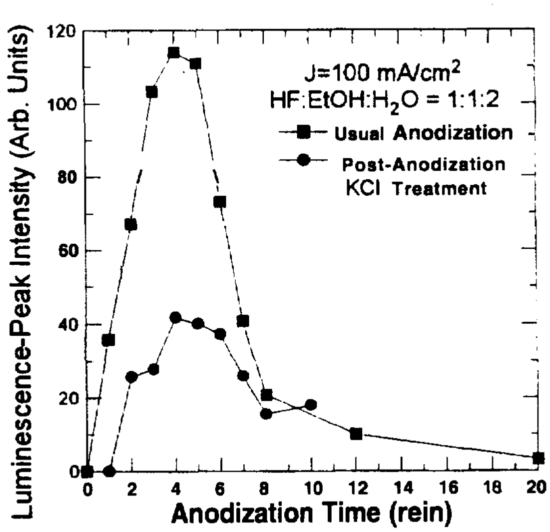


Fig. 8. Photoluminescence-peak intensity for an untreated and for a KCl treated sample vs. anodization time.

plotted the corresponding curve from Fig. 1. We notice that the KCl treatment does not shift the maximum to a different anodization time, however, it does lower the luminescence intensity. This latter is probably due to adsorption of some species, from the electrolyte. Thus we ascribe the shift of the anodization time for maximum luminescence, to possible surface damage due to the application of the voltage pulses.

CONCLUSION

P-type PS behaves quite differently from that of n-type PS. While the photoluminescence intensity exhibits a pronounced maximum with anodization time, the effective surface area and the surface state density appear to reach a more or less constant value as a function of anodization time. This is in contrast to n-type PS, where a close correlation between the effective surface area, the luminescence intensity and the surface-state density was found.[6] The discrepancy between the two results can be reconciled once we realize that in both cases the surface states are measured under accumulation conditions. The surface states involved are then those near the majority carrier band edge; namely, near the conduction band for the n-type material and near the valence band for the p-type material. Thus our present results suggest that the surface states near the valence-band edge are not involved in the luminescence process.

ACKNOWLEDGMENT

The work in Puerto Rico was supported by NASA grant No. NCCW-0088 and by NSF EPSCoR Grant No. OSR-94-52893, while the work in Israel was supported by the Ministry of Science, Israel, within the framework of the infrastructure-applied physics support project.

REFERENCES

1. L.T. Canham, Appl. Phys. Lett. 57, 1046(1990).
2. I. Amato, Science 252, 922 (1991).
3. A.G. Cullis and L.T. Canham, Nature 353,335 (1991).
4. For the recent developments in this field see: Z.C. Feng and R. Tsu, *Porous Silicon* (World Scientific, Singapore, 1994).
5. M. Wolovelsky, J. Levy, Y. Goldstein, A. Many, S.Z. Weisz and O. Resto, Surf Sci. 171,442 (1986).
6. S.Z. Weisz , J. Avalos, M. Gomez, A. Many, Y. Goldstein, and E. Savir in *Defect- and Impurity-Engineered Semiconductors and Devices*, edited by S. Ashok, J. Chevalier, I. Akasaki, N.M. Johnson, and B.L. Sopori (Mater. Res. Sot. Proc. 378, Pittsburgh, PA, 1995) pp. 899-904.
7. A. Many, Y. Goldstein, and N.B. Grover, *Semiconductor Surfaces* (North-Holland, Amsterdam, 1971).
8. S.Z. Weisz, J. Penalbert, A. Many, S. Trokman, and Y. Goldstein, J. Phys. Chem. Solids 51, 1067 (1990).

S134/39

URC97135

LARGE DEFORMATION TESTS ON HIGHLY FLEXIBLE STRUCTURES

Eric A. Wheater, P. Frank Psi, and Ahmad S. Naser
Structural Mechanics and Control Research Lab
North Carolina A&T State University
Greensboro, NC 27411

Abstract: This work is to establish a database of pre-buckling and post-buckling behavior of highly flexible structures and to verify the accuracy of an in-house nonlinear finite-element code GESA (Geometrically Exact Structural Analysis) for the analysis and design of deployable space structures. .4 test set-up that can test beams and frames subjected to different loading conditions, an accurate method of measuring large static deformations, and test results of a spring steel cantilever beam with a concentrated tip load will be presented. The experiment al results reveal the accuracy of numerical solutions from GESA. Difficulties in measuring large deformations involving large rotations and influence of initial imperfections will also be discussed.

1. Introduction

Highly Flexible Structures (HFSs) have been used in many mechanical systems, civil structures, and aerospace vehicles to reduce structural weight, satisfy space limitations, and/or provide special mechanisms. Moreover, because current NASA space missions require structures that have dimensions much greater than the shroud diameter of launch vehicles, deployable HFSs are extensively used for such structures. This extensive use of HFSs reveals the need for a design and analysis software package and a database system with guidelines for designing HFSs.

.4 new total-Lagrangian finite-element code GESA (Geometrically Exact Structural Analysis) has been under development for analyzing highly flexible structures. GESA is a displacement- based finite-element code written in MATLAB language and is based on newly developed theories for structures undergoing large displacements, large rotations, and finite strains [1-7]. The structural theories fully account for geometric nonlinearities due to large rotations, large in-plane strains of two-dimensional structures, large axial strains of one-dimensional structures, initial curvatures, and transverse shear deformations by using Jaumann stress and strain measures, an exact coordinate transformation, and a new concept of orthogonal virtual rotations. The Jaumann strains are derived using a new concept of local displacements without performing polar decomposition and they are proved to be a corotated objective measure. Because all possible initial curvatures are included in the strain-displacement equations, governing equations of plates and shells are unified and the strain-displacement relations can be used for most one- and two-dimensional structures. Only global translational degrees of freedom (DOFs) and their derivatives are used in the strain-displacement relations in GESA and no independent global or local rotational D OFs are defined. A corot ated point reference frame is defined using the symmetry of J aumann strains. Moreover, there is no need for transformations before updating strains, stresses,

and displacements. Cable, truss, bean-i, plate, and shell elements have been developed, and both isotropic and anisotropic materials are considered. However, experimental results are needed in order to verify the accuracy of GESA.

To verify GESA and to establish an adequate experimental database of pre-buckling and post-buckling characteristics of HFSs, we designed a test set-up and developed a method of measuring large deformations.

2. Test Set-Up

The flexibility of HFSs makes it difficult to measure displacements by conventional means. For example, any indicator that carries a small spring force does not supply accurate readings because the spring force results in a significant structural deformation. Electronic strain gages change the structural stiffness, and they cannot reveal rigid-body deformations because they only measure relative straining displacements. Air gages apply a force to the structure and deform the structure significantly. Laser gauging is an option because it is a non-contacting method, but the equipment is expensive and it is difficult to chase the measuring point when large rotations are involved. Several of these methods were evaluated based-on the goal of measuring large deformations of beams and frames, and some of them were actually tried. The final choice we made was to build an accurate reference metal frame that can be used to fix the test structure and to measure three-dimensional displacements from the frame. The metal frame we built is shown in Fig. 1. The two circular slots on the vertical wall can be used to adjust loading angles and hence different loading conditions can be tested.

Choosing an appropriate instrument to take the readings is another challenging issue. A vernier scale was chosen in the beginning, but its resolution is only .0078 *in* and it is difficult to judge the instruments proximity to the structure. If the instrument does not contact the structure and a sight of some kind is used, parallax is a problem. Our final decision was to use a dial caliper with a brass probe mounted on the tip, a series circuit, and an indicator lamp to ascertain the immediate contact of the probe and the structure. In measuring displacements, the probe is slowly moved toward the structure until the lamp lights. Figure 1 also shows the use of the dial caliper and the indicator lamp in measuring the transverse deflection of a horizontally mounted cantilever beam.

3. Measurement

A 15" x 2" x .02" spring steel cantilever beam was selected for tests. Since the Young's modulus E was not known, we measured the weight and volume to derive the mass density. Then we performed a linear vibration test by using a modal hammer, a proximity sensor, and a DP420 FFT analyzer to obtain a Frequency Response Function (FRF), as shown in Fig. 2. Using the derived mass density, the first few natural frequencies from the measured FRF (see Fig. 3), and the formula of linear natural frequencies of a cantilever beam [8], we obtain that E = 2.84 x $10^7 psi$ and the mass density is $=.271 lb/in^3$.

Figure 1: Test set-up.

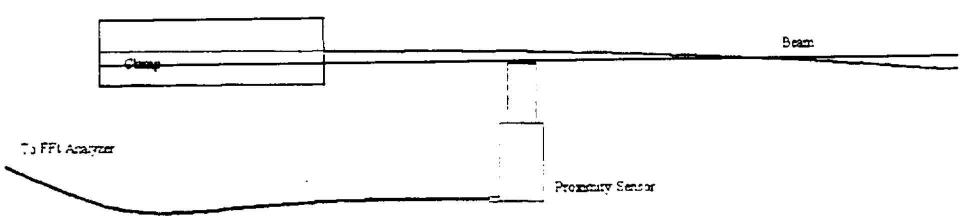


Figure 2: A linear vibration test for obtaining frequency response functions.

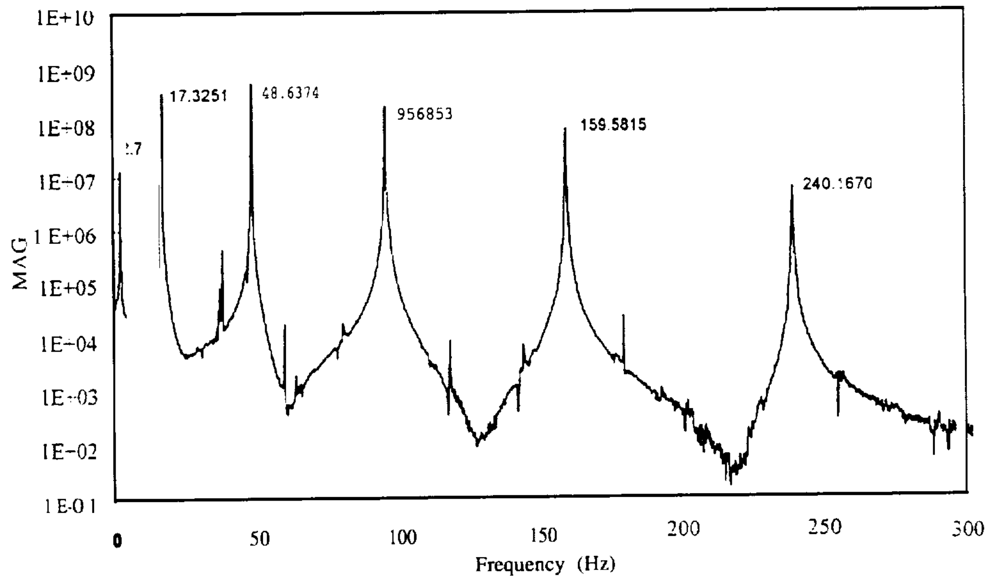


Figure 3: The measured frequency response function.

Displacements were measured at fifteen points in 1" increments along the beam, as shown in Fig. 4. Measurements were taken from the metal frame to each point on the beam using the modified dial caliper. Initial displacements due to imperfection in the sample were measured and recorded. Then load was applied in increments of 10 grams up to a maximum of 100 grams. For each load, transverse and longitudinal displacements of each point were measured, and three different measurements were taken and the results were averaged. The displacements due to imperfection were subtracted from the averaged displacements of each point under each load.

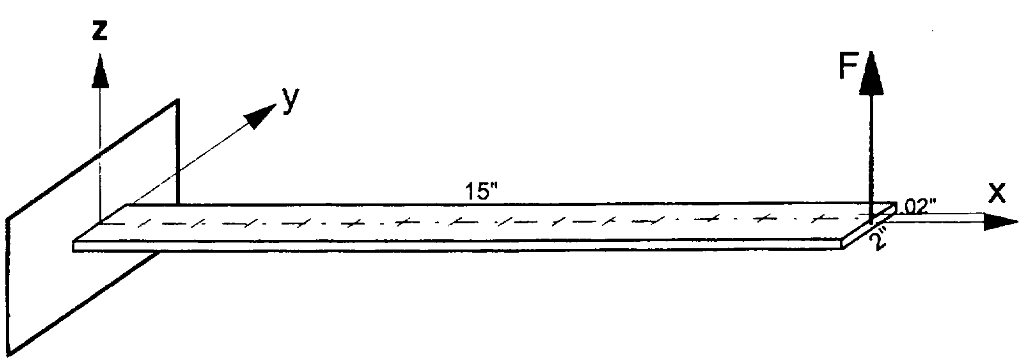


Figure 4: The measuring points on the cantilever beam.

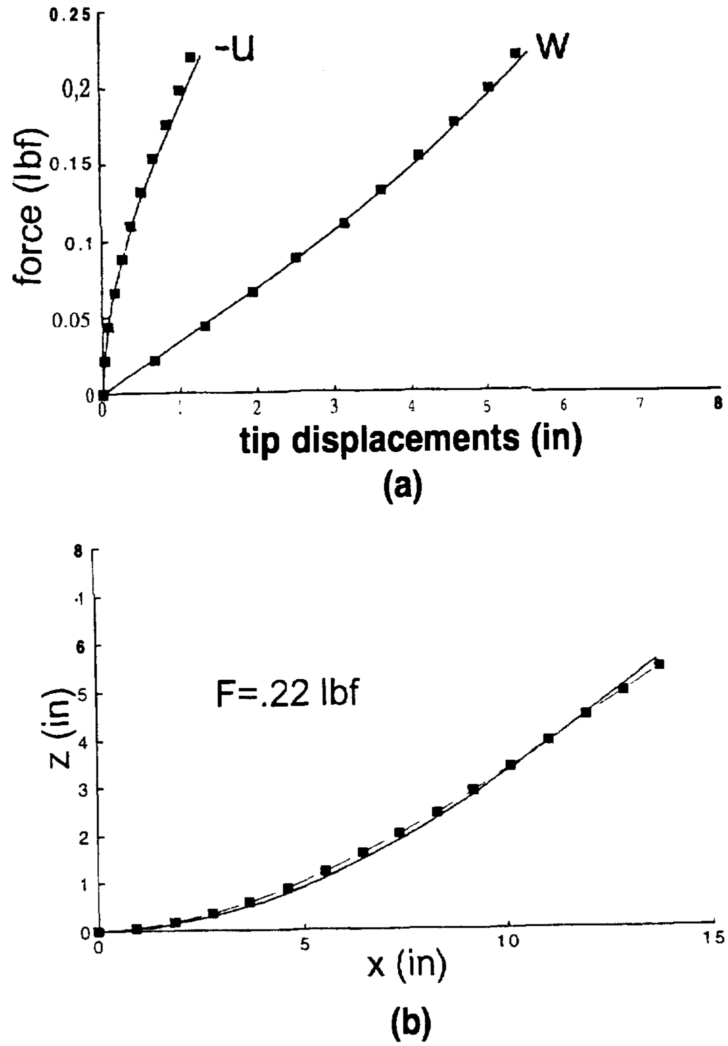


Figure 5: Test results: (a) the transverse and longitudinal displacements of the tip with respect to different loading values, and (b) the deformed geometry when $F = .22\ lbf$.

4. Experimental Results

Figure 5 shows a typical test result, where the small solid squares are the experimental data and solid lines are numerical solutions obtained from GESA by using eleven equal beam elements. influence of transverse shear deformations was investigated in the finite element analysis. It was found that transverse shear deformations are negligible for this thin flexible beam.

5. Discussion

The obtained experimental results show that GESA is accurate in predicting the large static deformations of the specific flexible beam. However, to fully verify GESA tests on other different structures are necessary, especially flexible composite structures with elastic bending-torsion and extension-torsion couplings. We are in the progress of testing L-shaped frames and circular rings. Stability, deployability, and post-buckling characteristics of such

HFSs will be studied in-depth. Future work also includes testing nonlinear dynamics of HFSs to verify the accuracy of dynamic solutions from GESA.

References

1. Pai, P. F. and Palazotto, .4. N., "Nonlinear Displacement-Based Finite-Element Analyses of Composite Shells — A New Total Lagrangian Formulation," *Int. J. Solids and* Structures 32, 3047-3073, 1995.

2. Psi, P. F., "A New Look at Shear Correction Factors and Warping Functions of Anisotropic Laminates," *Int. J. Solids and* Structures *32,* 2295--2313, 1995.

3. Psi, P. F. and Palazotto, A. N., "Polar Decomposition Theory in Nonlinear Analyses of Solids and Structures," *J.* Engineering Mechanics *121(4),* 568-581, 1995.

4. Psi, P. F. and Nayfeh, A. H., "A New Method for the Modeling of Geometric Nonlinearities in Structures," Computers & Structures 53, 877–895, 1994.

5. Pai, P. l?. and Nayfeh, A. H., "A Unified Nonlinear Formulation for Plate and Shell Theories," *Nonlinear* Dynamics *6,* 459–500, 1994.

6. Pai, P. F. and Nayfeh, A.. H., "A Fully Nonlinear Theory of Curved and Twisted Composit e Rotor Blades Accounting for Warpings and Three-Dimensional Stress Effects ," *Int.* J. Solids *and* Structures 31, 1309-1340, 1994.

7. Psi, P. F. and Nayfeh, A. H., "Fully Nonlinear Model of Cables," *AIAA Journal* 30, 2993-2996, 1992.

8. Inman, I). J., *Engineering* Vibration. Prentice-Hall, New Jersey, 1994.

N 1 1 - 1 4 135/33

URC97136

Characterizing Surfaces of the Wide Bandgap Semiconductor Ilmenite with Scanning Probe Microcopies

R. Wilkins
Center for Applied Radiation Research &
Department of Electrical Engineering
Prairie View A&M University
Prairie View, Texas 77446

Kirk St. A. Powell
Department of Electrical Engineering
Prairie View A&M University
Prairie View, Texas 77446

Abstract:

Ilmenite ($FeTiO_3$) is a wide bandgap semiconductor with an energy gap of about 2.5eV. Initial radiation studies indicate that ilmenite has properties suited for radiation tolerant applications, as well as a variety of other electronic applications. Two scanning probe microscopy methods have been used to characterize the surface of samples taken from Czochralski grown single crystals. The two methods, atomic force microscopy (AFM) and scanning tunneling microscopy (STM), are based on different physical principles and therefore provide different information about the samples. AFM provides a direct, three-dimensional image of the surface of the samples, while STM give a convolution of topographic and electronic properties of the surface. We will discuss the differences between the methods and present preliminary data of each method for ilmenite samples.

Introduction:

The progress of technology creates new demands for discovering and developing new electronic materials with increasingly improved properties. Of particular interest to NASA are those materials that allow both space and commercial applications. Wide bandgap semiconductors like ilmenite have properties that make them attractive for a variety of applications in both the terrestrial and space realms. Wide bandgap semiconductors have potential use in high power, high frequency, high temperature microelectronic and optoelectronic (including photovoltaic) devices which are also resistant to radiation damage.

To exploit an electronic material for these types of applications, the properties of the surface are essential for developing optimum processing procedures for device fabrication. Two new surface analysis tools proved useful in microelectronic fabrication: Atomic Force Microscopy (AFM) and Scanning Tunneling Microscopy (STM). These microcopies have resolution at the atomic level, yet can be used to examine device geometries in ultra-large scale integration ULSI circuits[1].

This paper will discuss ilmenite and some of its potential applications. We will then briefly introduce the AFM and STM techniques, discussing the information these methods provide and their limitations. We then present and discuss preliminary AFM and STM data characterizing ilmenite surfaces.

Ilmenite:

Large, high quality Czochralski grown single crystals have been produced to evaluate the materials performance for a number of applications[2]. In particular, the photovoltaic, optoelectronic, and the high radiation resistant characteristics of ilmenite seem to offer superior performance compared to traditional semiconductors. For instance, silicon solar cell has been the mainstay power conversion source for the U.S. space program. However, the conversion efficiency of Si seems to be bound at an upper limit of ~30%, and this efficiency degrades with time[3]. In addition, a wide bandgap semiconductor such as ilmenite can take advantage of the lower wavelength end of the solar spectrum. Other potential space application benefits of ilmenite include:

- Weight Savings: Extensive cooling equipment is required for Si based electronic on spacecraft. Much of this equipment could be eliminated using a large band gap material such as ilmenite.
- Radiation Resistance: Electronics in space operate without the benefit of the atmosphere to shield them from the relatively high ambient radiation environment. Ilmenite has appears to have radiation resistant properties[4].

Potential commercial applications of ilmenite include:

- Optoelectronic Devices: Ilmenite appears to have a direct bandgap making it a candidate for blue-green lasers diodes.
- Thermoelectric Coolers: Initial measurements of Seebeck coefficients appear to make ilmenite attractive in refrigeration and cooling systems.
- Heterostructures of Ilmenite: The interesting annealing properties of ilmenite (which effect the oxidation state of the iron) may provide the means for creating "seamless" interfaces between areas of n- and p-type material, perhaps leading to tunable bandgaps.

Ilmenite can be grown both an p-type and an n-type semiconductor making bipolar junction devices consisting of alternate n- and p-type materials possible. Some of ilmenite's basic properties are summed up in Table I.

Physical Property	Ilmenite $FeTiO_3$
Bandgap	2.58 eV direct(?)
Crystal Structure	Corundum (Al_2O_3)
Unit Cell	Hexagonal
a	5.09 A
c	14.09 A
c/a	2.77
Resistivity	1.45 Ωm
Hall coefficient	$0.26 * 10^{-5} m^3/C$
Carrier Concentration	$2.6 * 10^{+24} m^{-3}$
Density	4.83 gm/cm^3
p-type semiconductor	pure ilmenite
n-type semiconductor	in solution with a-Fe_2O_3
Melting point	1410°C

TABLE I: Properties of Ilmenite

Scanning Probe Microscopy:

Scanning probe microscopy techniques have gone from Nobel Prize winning discovery to production line characterization [5] in a little over a decade. Starting with the scanning tunneling microscope (STM) in 1982 [6], many different SPM techniques have been described in the literature [7]. These microscopes allow surface characterization of both conductors and insulators with sub-nanometer resolution.

Atomic Force Microscopy: Figure 1 shows a schematic of an AFM configuration. The AFM monitors the minute forces between a cantilevered probe tip and a sample surface. When the tip is scanned parallel to the sample, surface topography will tend to deflect the cantilever. The deflection is detected by reflected laser light via a photodetector. A feedback circuit monitoring this deflection controls a piezoelectric element (not shown in Figure 1) which moves the sample to counter the cantilever deflection. In effect, this keeps the force between tip and sample constant. The signal to the piezo is monitored and a topographic image is constructed from this signal as the probe is raster scanned (also by piezo elements not shown) across the sample. The typical value of force involved with AFM is 10^{-9}N, and examples of true atomic resolution have been obtained[8].

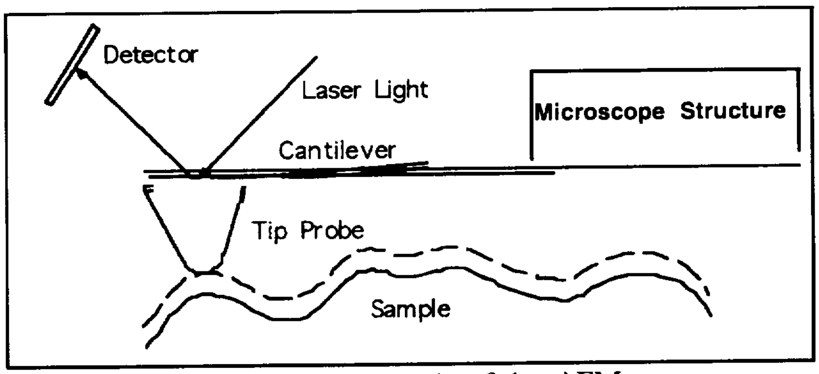


Figure 1: Schematic of the AFM.

One of the advantages of the AFM is that the technique works on insulators as well as conductors and is capable of routine nanometer resolution in ambient conditions. Moving the tip perpendicularly with respect to the sample surface provides data on the force of interaction between tip and sample and can, for example, provide information on the force of adhesion to a sample surface. The lateral force mode measures the forces on the tip cantilever parallel to the surface. This is a measure of the frictional forces between a probe tip and sample surface, making the AFM an ideal tool for tribological [9, 10] and wear [11] studies.

Scanning Tunneling Microscopy: Quantum mechanical tunneling of electrons through a classically forbidden energy barrier is exponentially sensitive to the barrier width, giving the scanning tunneling microscope (STM) unique three-dimensional atomic resolution [12]. The STM is schematically illustrated in Figure 2. A metal tip is brought within about 1 nm of the surface of a conductive sample. An applied voltage between the tip and sample establishes a small (on the order on nanoamps) but measurable tunnel current between the tip and sample. A feedback loop maintains a constant tunnel current as the tip is raster scanned *along* the surface. By recording the response of the feedback loop, an image of the surface is obtained. The tip-sample separation and raster scan are controlled by piezoelectric elements.

The interpretation of STM images is not always straightforward [13,14]. The tunnel current measured by the STM is not only proportional to the barrier width (given by the tip-sample separation) but also the local density of electronic states at the surface. Therefore, an STM image does not necessarily reflect only the topography of a surface. As AFM images should represent the actual topography of a surface (ignoring tip geometry artifacts), comparisons between AFM images and STM images have been made to check if a STM image has "electronic effect s" involved. An excellent example of this is the study of some high-temperature superconducting film[15]. Some of these sample (depending on their growth conditions) exhibited a surface made up of spiral structures superimposed on one another. It was determined by AFM images of the films that the relatively higher resolution STM images were indeed representative of the surface topography.

On a conducting sample, STM will generally provide superior resolution as compared to the AFM. In addition, the fact that the tunnel current is dependent on the local electronic density of states can be used to measure the spectra of electronic states near the Fermi level. In this mode, the tip is held at a point on the sample and the feedback loop is momentarily disabled. During this time, the voltage between the tip and sample is varied and the current response measured, thus measuring the current-voltage (I-V) characteristics of the tunnel junction between tip and sample. This is repeated to form an array of I-V characteristics over the sample surface; this technique is called scanning tunneling spectroscopy.

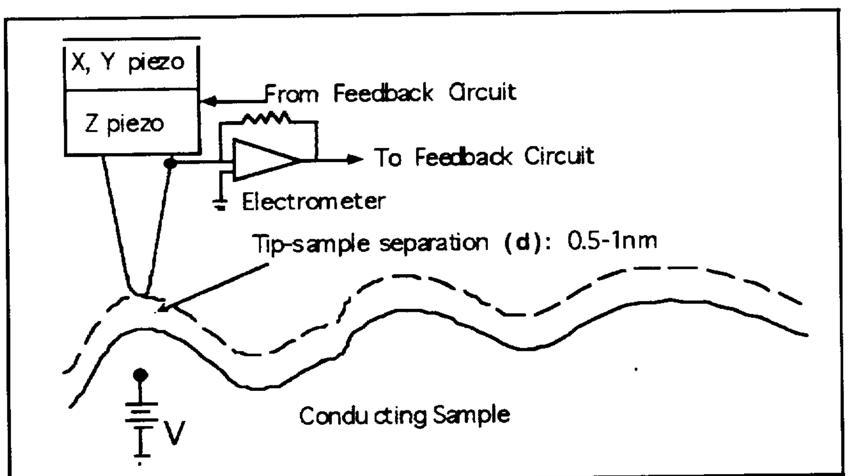


Figure 2: Schematic of the STM

Results and Discussion:

Preliminary data is shown below; Figure 3 shows an AFM image and Figure 4 shows a STM images of ilmenite surfaces. The AFM image is a 13 .4umx13 .4um scan. The sample was microscopically flat but the microscopic image reveals a fairly rough surface. The deep grooves evident in the image are likely due to the wafer polishing process. The overall surface roughness (in the large area on the left side of the image) is about 20-30nm. The STM image is a 1.5umx1.5um scan taken at a tunnel current of about 200pA at a junction bias of 3V. (Wide bandgap semiconductors are difficult to image with an STM because of their relatively high resistivity.) Microscopically the sample had few flat regions and appeared layered. The image was taken on a large (~3mm^2) flat area; the microstructure of this surface was characterized by parallel streaks. It is possible that the streaks in the STM image represent the layered crystallite structure (ilmenite has hexagonal crystal structure). The streaks in the STM image do not represent polishing artifact because this was not a polished sample. Being able to take STM data

on ilmenite is encouraging because it holds the prospect of studying the density of states near the Fermi level with tunneling spectroscopy. The two images indicate about the same degree of overall surface roughness (20-30nm) for the "flat" region of the samples. However, given their relative scales, little more can be conclusively said about the two images. Work is in progress to further characterize ilmenite with these technique.

Conclusions:

- Ilmenite is a wide bandgap semiconductor with a variety of potential applications, including radiation resistant electronics.
- Scanning probe microcopies provide important information about the topography (AFM, STM) and electronic properties of a surface (STM).
- STM studies electronic properties of a conductive surface, while an AFM gives topographic information on both insulators and conductors.
- STM is possible on ilmenite, and thus tunneling spectroscopy.
- Additional SPM work is needed on ilmenite.

Acknowledgments:
This work is partially supported by the NASA/PVAMU Center for Applied Radiation Research and the Texas Advanced Technology Program. Data for Figure 3 was obtained by one of us (K. P.) at the Electron Microscopy Center of Texas A&M University.

References:

[1] G. Vachet and M. Young, *Solid State Technology 38,57* (December 1995).
[2] A. A. Kumar, T. N. Fogarty, R. K. Pandey, and R. Wilkins, *Proceedings of the Dual-Use Space Technology Transfer Conference*, NASA Conference Publication 3263, p. 347 (1994).
[3] D. J. Flood, *Chem. Engin. Prog., 85 (4), 62 (1989).*
[4] J. N. Mitchell, Center for Material Science, Summer Research Group Seminar, July 1996.
[5] "A Kinder, Gentler Chip Inspection", *Science* 258, 1575 (1992).
[6] G. Binnig, H. Roher, Ch. Gerber, and E. Weibel, *Phys. Rev. Lett. 49,57 (1982).*
[7] H. Kumar Wickramasinghe, in *Scanning Tunneling Microscopy*, Joseph A. Stroscio and William J. Kaiser, eds. (San Diego, Academic Press, 1993) p. 77.
[8] C. F. Quate, in *Surface Science, The First Thirty Years*, C. B. Duke, ed. (Amsterdam, North-Holland, 1994) pg 980.
[9] Ju-Ai Ruan and Bharat Bhushan, *J. of Tribology 116,378 (1994).*
[1 O] I. L. Singer, *J. Vat. Sci. Technol. A* **12,2605 (1994).**
[11] M. P. Everson, A. K. Gangopadhyay, R. C. Jaklevic and D. Scholl in *Forces in Scanned Probe Methods*, Guntherodt et. rd., eds., (Kluwer, Dordrecht, 1994).
[12] R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy* (Cambridge University Press, 1994), Chapter 1.
[13] R. M. Feenstra, J. A. Stroscio, J. Tersoff, and A. P. Fein, Phys. Rev. Lett. 58, 1192 (1987).
[14] G. Lengel, R. Wilkins, G, Brown, M. Weimer, J. Gryko, and R. E. Allen, *Phys. Rev. Lett 72,836 (1994).*
[15] Ian D. Raistrick and Marilyn Hawley, in *Interfaces in High-T_C Superconducting Systems*, Subhash L. Shinde and David A. Rudman, eds. (New York, Springer-Verlag, 1994), p. 28.

Page intentionally left blank

/ 1 ... 313 145

URC97137

Calibration of the QCM/SAW Cascade Impactor For Measurement of Ozone

Cassandra K. Williams, C. B. Peterson, V. R. Morris

Department of Chemistry
Center for the Study of Terrestrial and Extraterrestrial Atmospheres
Howard University
Washington, D.C. 20059

Abstract

The Quartz Crystal Microbalance Surface Acoustic Wave (QCM/SAW) cascade impactor is an instrument designed to collect size-fractionated distributions of aerosols on a series of quartz crystals and employ SAW devices coated with chemical sensors for gas detection. We are calibrating the cascade impactor in our laboratory for future deployment for in-situ experiments to measure ozone. Experiments have been performed to characterize the QCM and SAW mass loading, saturation limits, mass frequency relationships, and sensitivity. The characteristics of mass loading, saturation limits, mass-frequency relationships, sensitivity, and the loss of ozone on different materials have been quantified.

Introduction

The QCM/SAW flight cascade impactor is an analytical instrument composed of twelve contiguous stages that can measure eight size-segregated fractions of aerosols and detect trace gases down to sub parts per billion levels simultaneously. It is a compact, lightweight, and sturdy device that is one of the principal instruments used for the characterization of atmospheric aerosols. Since 1979, QCM crystals have been used for collecting size-segregated samples of stratospheric aerosols. [1-2] The Quartz Crystal Microbalance cascade imp actor has been used for characterization of volcanic eruption materials on the climatic effects of the stratosphere. [3]

The quartz crystal microbalance is an extremely sensitive sensor capable of measuring mass changes in the nanogram range. For a 10 MHz AT-cut crystal, the sensitivity is approximately 10^9 Hz/g. [4] A schematic of the QCM crystal shown in Figure 1. In this application, an electric voltage is applied to the piezoelectric quartz crystal (PQC) which induces a shear stress resulting in an oscillation through the body of the 10 MHz QCM crystal confined to the region between the electrodes. In the SAW device, a 200 MHz acoustic wave propagating along the surface of the crystal betwen a pair of interdigitated electrodes is induced by the applied voltage. Each individual stage contains two crystals where the upper crystal is used for aerosol or gas impaction, and the lower reference crystal compensates for frequency shifts due to temperature changes. The change in beat frequency is monitored to give near-real time measurements of the mass of deposited aerosol.

The SAW crystals have the ability to measure various gases and the mass of very small particles with a higher sensitivity. The primary application of the SAW crystal in mass-measuring

instruments has been in the area of chemical sensing, in which a reactive coating applied to the surface of a SAW crystal responds specifically to a given gas through a mass change which in turn results in a frequency change. [5]

The focus of the calibration of the QCM/SAW was to determine the sensitivity, accuracy, and detection limits of the QCM and SAW crystals for ozone measurement. It was also necessary to quantify the loss of ozone through the flight stack so our data from field experiments could be accurately analyzed. The initial motivation for this work was the need for trace gas measurements to accompany aerosol measurements during stratospheric flight experiments.

Experimental Methodology

The experimental conditions for our calibration experiments simulated the sampling conditions at 18-21 km which is the altitude at which the field experiments would be performed. The calibration was performed for ozone concentrations ranging from 1-500 ppb. Two types of experiments have 'been performed for ozone; mass sensitivity/ mass loading calibration experiments and ozone loss experiments.

Preparation for a typical laboratory experiment consists of coating a crystal with a dilute (O. 1- 1.0% by mass) solution of polybutadiene in toluene. The crystal is then placed in the QCM/SAW stack where a stream of ozone is passed over the crystal for a certain length of time. A Thermoenvironmental Model 49 PS, Ozone Calibrator is used to supply the ozone to the QCM/SAW. In the ozone loss experiments, a Thermoenvironmental Model 49 C Ozone Analyzer was attached to the outlet port of the QCM/SAW stack to quantify the loss of ozone through the instrument and through transfer tubes of different materials. Both the ozone calibrator and analyzer are certified and traceable to NIST standards. The flow rate is controlled by a micrometering valve and monitored by a Matheson Model FM 1050 flow meter and exhausted with the an oiless sampling pump. Before each experiment, the beat frequency is stabilized by a nitrogen gas stream (99.995%, Potomac Airgas) directed across each of the crystals consecutively.

Mass loading is defined as the accumulation of mass onto a surface by chemical reaction, chemisorption, or adsorption or by diffusion into the bulk of the surface coating. The two primary characteristics of the mass loading properties of any crystal microbalance system are mass sensitivity and saturation limits. From the analysis of the mass loading response curves, an average mass loading is obtained for each set of experimental conditions. Total mass loadings from a series of response curves at a given flow rate and temperature produce a calibration curve. Mass-frequency relationships can be determined from this analysis. Over the course of the time that ozone is pumped across the crystal an irreversible chemical reaction occurs between ozone and the chemical coating (ozonolysis). The product of this reaction remains on the crystal causing a change in the beat frequency of the dual oscillator system. This change in frequency is monitored electronically and stored on a computer for later analysis. Mass loading characteristics for SAW crystal at 50 ppb shown in Figure 2.

Calibration experiments for the ozone loss incorporated use of the photometric ozone analyzer to measure the losses of ozone for concentrations ranging from 25-500 ppb. Two separate sets of experiments were performed to characterize the ozone loss: loss of ozone through different materials and loss of ozone through the flight stack. During these experiments, the Ozone Analyzer was attached on 1) the outlet port of the QCM/SAW stack or 2) the Ozone Calibrator connected by monel, aluminum, stainless steel, copper, and teflon in order to further characterize the possible loss processes for ozone on different materials. After stabilizing the calibrator and analyzer, ozone was passed through the tubings for one to four hours. The two lengths used were 17 $^1/_4$ in. and 58 $^1/_2$ in. with surface areas of 13.54 in^2 and 49.9 in^2 respectively. The studies helped elucidate the most efficient design for gas flow into the gas detecting stages.

In the second set of experiments, the ozone was pumped through the entire flight stack to determine the loss of ozone through the stack. The experiment ran for five hours to correspond to the flight time of an actual field experiment. Teflon was used as the tubing to deliver the ozone to the QCM/SAW stack and ozone analyzer.

Discussion and Future Directions

The primary findings obtained from this study were: 1) that the 200 MHz SAW crystal is problematic and imprecisely measures mass loading by trace gases; 2) the 10 MHz QCM crystals are sensitive enough for the measurement of stratospheric, ozone; and 3) the loss of ozone through the flight stack has been quantified, and teflon has been determined to be the best tubing for minimal ozone loss to the flight stack.

The SAW stages were originally used for gas detection. Hence, they were also first used for the calibration of ozone. Maintenance of the crystal proved to be impossible for long time or repetitive usage. The response of the SAW crystal degrades noticeably on exposure to air. The recommended cleaning procedure for the SAW crystals was ineffective and accelerated the degradation. of the SAW crystal response. The crosslinked reaction product from ozonolysis could not be removed by washing with the routine organic solvents (toluene, benzene, and methanol), and even the plasma cleaning method was insufficient to completely clean the oxygenated residue left after ozone had reacted with polybutadiene. After plasma exposure times as short as two minutes, there was observation of both electrode and crystal surface degradation. Longer or accumulative exposures resulted in even more severe damage to the SAW crystals. Observations indicate that the SAW crystals, if not pristine, are not generally reusable and further, that plasma cleaning and reuse will lead to irreversible degradation and potentially unreliable results. Due to these problems with SAW crystals, the QCM crystals became the prime candidates for the calibration of ozone on the instrument. We have found that the 10 MHz QCM crystals were sensitive enough for ozone detection down to the low ppb concentration ranges.

The QCM crystals were much more cost effective than the SAW crystals even though the cleaning method proved to be destructive to these crystals as well. To obtain accurate data, crystals had to be discarded and replaced with new crystals. Several cleaning methods were used to clean the QCM crystals after exposure. The QCM crystals were observed to have a longer lifetime and could be reused but eventually experienced performance degradation as well.

Typically, the QCM crystals were cleaned by rinsing only with the solvent of the polymer coating and then brushing it slightly with a Q-tip.

Polybutadiene was found to be selective and specific for the detection of O_3 to concentrations in the low ppbv range. This coating yielded a linear response in beat frequency as a function of the concentration under standard operating conditions. Figure 3 illustrates the mass loading response for a SAW crystal at 50 ppb. The lower bound to the sensitivity of the SAW crystals has been found to be 0.05 ng. This corresponds to ambient concentrations in the low ppb range. The mass-frequency constant is 3.0×10^{10} Hz/ng.

Ozone loss studies provided essential information to optimize the QCM/SAW stack with teflon being the best tubing for minimal ozone loss. Statistical analysis of the ozone loss data for teflon obtained for ozone concentrations between 25-500 ppb yielded an average loss of approximately 3%. This is in good agreement with previous ozone loss studies of the teflon coated Dasibi ozone monitors. [6] The ozone loss on stainless steel over the same range of concentrations was approximately 5%.

Acknowledgments

This research was supported by NASA funding under grant number NAGW-2950. The authors acknowledge other members of the Center for the Study of Terrestrial and Extraterrestrial Atmospheres director, Dr. Arthur N. Thorpe, Hideo Okabe, and Silas Anamelechi.

References

1. R.L. Chuan; D.C. Woods; M.P. McCormick, Science. 211, 830-832 (1981)

2. R.L. Chuan and D.C. Woods, Geophys. Res. Letts. 11,553 (1984)

3. D.C. Woods and R.L. Chuan, Characteristics of Aerosols in the Lower Stratosphere in Aerosols and Climate, P.V. Hobbs and M.P. McCormick (Eds.) A. Deepak Publishing

4. G.Z. Sauerbrey, Phys. 155,206 (1959)

5. R.L. Chaun and W.D. Bowers, Rev. Sci. Instrum. 60,1297 (1989)

6. J.E. Ainsworth, J.R. Hagemeyer, and E.I. Reed, Geophys. Res. Letts. 8, 1071 (1981)

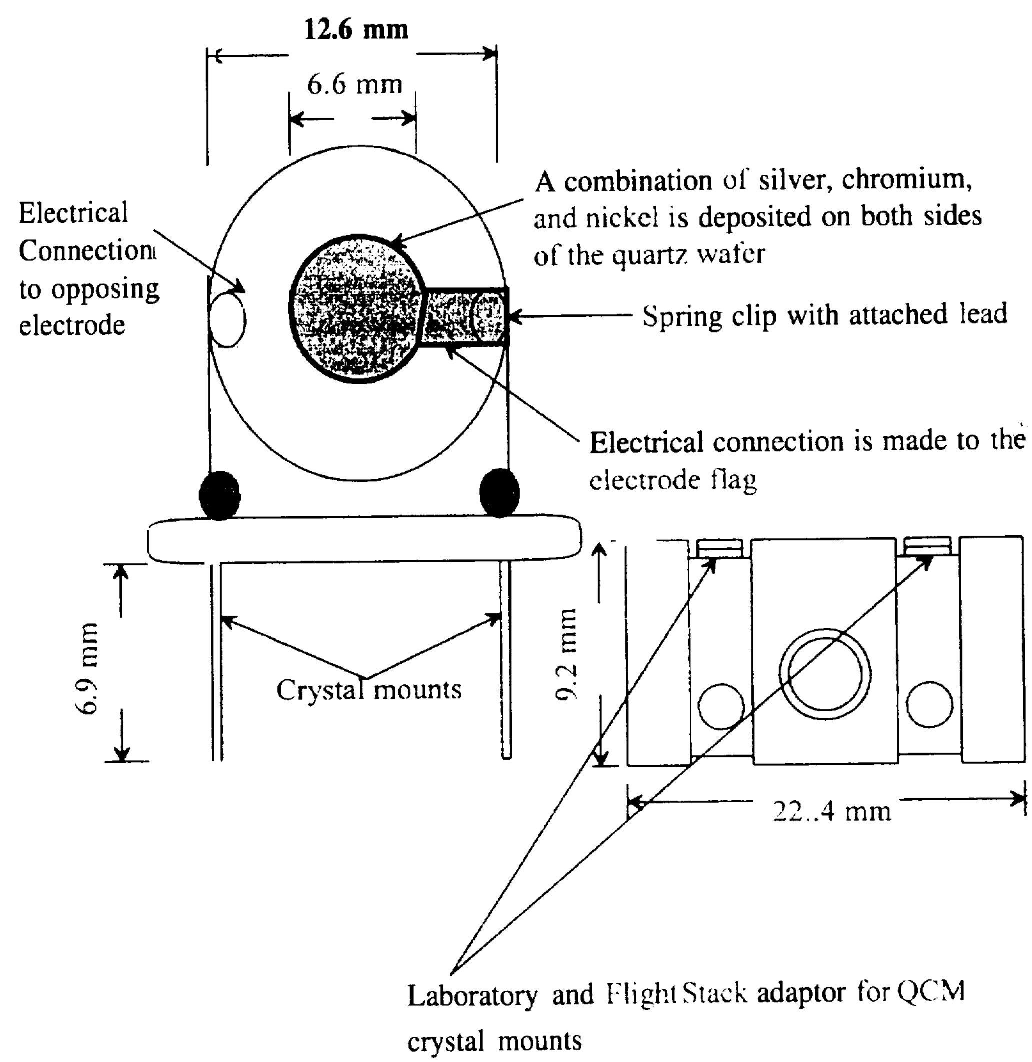


Figure1. Schematic representation of QCM crystal

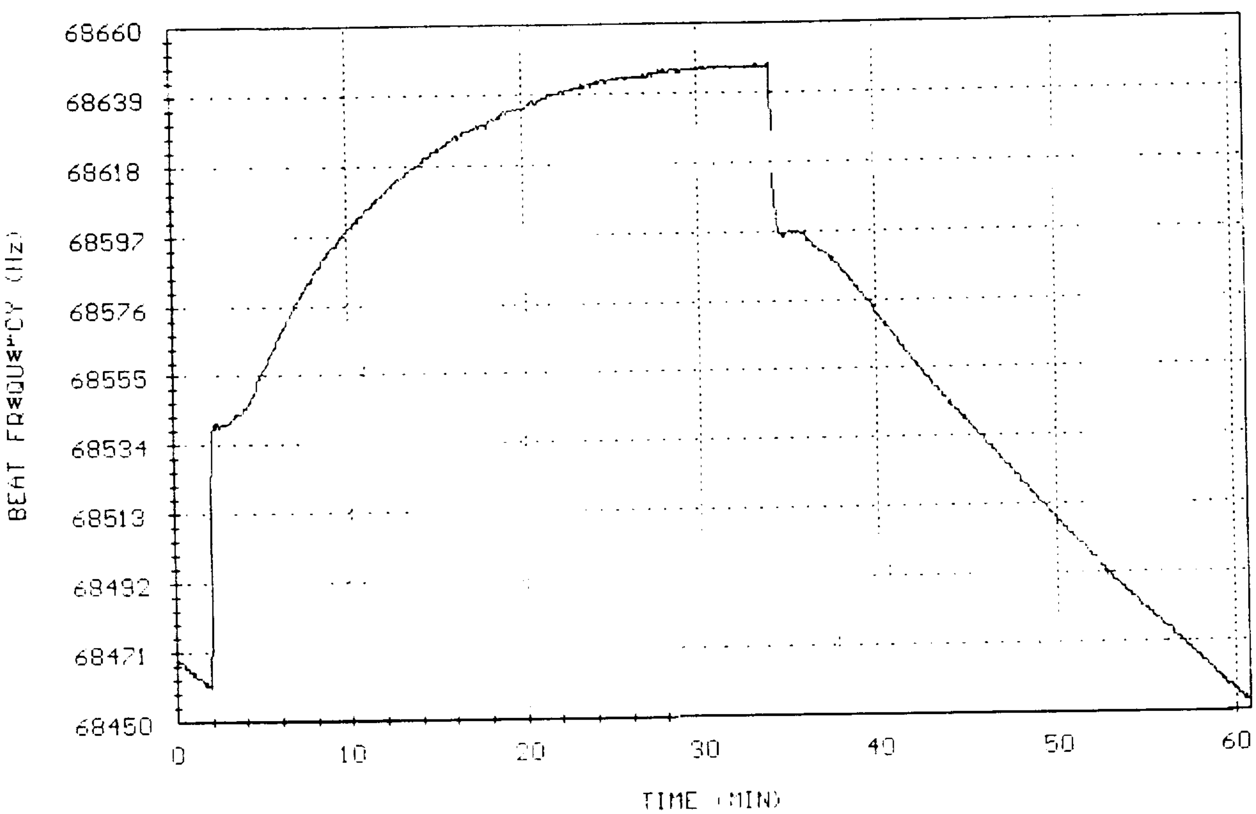


Figure 2. Response Curve of Polybutadiene with 50 ppb of ozone on SAW crystal

URC97138

Modeling and Simulation of Photo-CVD Reactors

Melanie Williams, Sy-Chyi Lin and Jorge Gabitto
Chemical Engineering Department
Prairie View A&M University

5/37/-7

ABSTRACT

Processes for deposition of thin films at low temperatures are needed in the microelectronics industry due to the increasing miniaturization of electronic devices. Laser-induced chemical vapor deposition (LICVD) has been used to deposit Si and Ge films from germane and silane. However, despite the importance of these processes the mechanisms that lead to the deposition are not well understood yet. In order to obtain better design of LICVD reactors a mathematical model is developed. The model is based on standard literature assumptions and experimental information. The active chemical species within the gas phase are selected using experimental information. Boundary layer equations are used to calculate the ftemperature and flow profiles. Mass balances are derived for all the active species. A simplified surface kinetics is proposed. Finite-difference methods are employed to solve for the two-dimensional fluid active species concentrations distributions. Reaction rates under different operating conditions are calculated.

INTRODUCTION

The development of techniques for low-temperature semiconductor film growth is currently an area of intense research interest for microelectronic applications including thin-film transistors [1] and large area devices such as liquid-crystal displays and solar cells [2]. In LICVD from germane and silane the initial photofragments are germane photolytic products formed very near the substrate surface at low pressures. [3]. The input gas is irradiated by a high energy laser beam. The gas phase after irradiation contains neutral radicals, the presence of which enhances the growth rate, and makes the *in situ* incorporation of dopants into the lattice feasible at low temperatures. Therefore, LICVD addresses the problems of conventional CVD while retaining the advantages.

A knowledge of the nature of the species which form the dominant precursors for film growth under a given set of depositions conditions, together with the mechanisms by which these species interact with the substrate and growing film is of fundamental importance for understanding and modeling film growth. Recently, some researchers [4] studying Si deposition from disilane using LICVD found that the predominant precursor is a "close shell" species, H_2SiSiH_2. A deposition yield of Si atoms from ArF excimer laser photolysis of disilane of approximately 20 % was reported.

The main goal of this work is to develop a comprehensive model that addresses the complex phenomena that determine the deposition of *a*-Si hydrogenated thin films *in* continuous parallel electrodes LICVD reactors. This approach will yield valuable knowledge on thin film technology.

MODEL DEVELOPMENT

A CVD deposition reaction always comprises three fundamental steps, gas phase reactions, transport towards the solid surface, and heterogeneous reaction on the substrate surface. Precursors generation in gas phase, mass transport phenomena from the plasma to the surface, and surface kinetics are the steps controlling the growth rate. The determination of the actual mechanism involves knowledge of the intermediate species formed in gas phase and/or the substrate surface,

The model proposed here rest heavily on the experimental results determined by Fowler et al [4]. They deposited Si thin films using an ArF excimer laser parallel to the substrate surface in a cylindrical reactor. The used a single wafer reactor with a laser beam parallel to the surface. The gas feed is also parallel to the wafer surface, but flows perpendicular to the laser beam (Figure 1), Fowler et al [4] reported that the use of an excimer laser as the energy source presents distinct

advantages over other photo-CVD methods. A photon energy of $h\nu = 6.4$ eV dissociates Si_2H_6 by single photon absorption, producing Si containing radicals that lead to film growth.

In order to estimate the fraction of photolysis products that can reach the film by mass transport from the beam-excited region, the quantum yields (ϕ) for formation of the various photolysis products and the gas phase kinetics of the individual products must be considered. The number of each photofragment produced by irradiation is given by the product of the quantum yield for the different species and the number of excited disilane ($Si_2H_6^*$) molecules, N. The dissociated quantum yield in Si_2H_6 photolysis at 193 nm has been measured to be 0.7 ± 0.1. The additional 30 % of the $Si_2H_6^*$ is thought to decay radiatively or be stabilized collisionally [4]. However, it is unlikely that significant gas phase nucleation is occuring under the operating conditions reported by Fowler et al. [4]. THe number of Si_2H_6 molecules under the substrate that absorb a photon can be estimated by using the photoabsoprtion cross-section of Si_2H_6 at 193 nm, $\sigma = (3.4 \pm 0.3)$ 10-18 cm2 [4], the Si_2H_6 concentration, n, and the path length of the beam inside the chamber. Fowler et al. [4] derived the following equation that gives the intensity absorbed, I_a, by the Si_2H_6 molecules,

$$I_a = I \exp[-\sigma n L]\,(1 - \exp[-\sigma n L\,]) \qquad (1),$$

where L and d are shown in Fig. 1, and I is the measured laser intensity in J/cm2 pulse that enters into the chamber through the laser inlet window. The number of photons absorbed per pulse, N, is I_a divided by the photon energy, $h\nu$, and multiplied by the beam cross sectional area, WH. Under optically thin absorption conditions, which occur at low Si_2H_6 partial pressures we get,

$$N = I\,\sigma\,n\,\frac{V}{h\nu} \qquad (2),$$

where V is the volume of the beam under the substrate, which is called as the active volume. The number, N, is also the number of excited Si_2H_6 molecules, $Si_2H_6^*$, in the active volume.

The following main reactions in gas phase is proposed based on the experimental information reported by Fowler et al. [4].

$$Si_2H_6\,(g) + h\nu \xrightarrow{k_1} SiH_2^*(g) + SiH_4(g) \qquad (3)$$

$$Si_2H_6\,(g) + h\nu \xrightarrow{k_2} H_3SiSiH^*\,(g) + H_2\,(g) \qquad (4)$$

$$Si_2H_6\,(g) + h\nu \xrightarrow{k_3} 2\,SiH_3^*\,(g) \qquad (5)$$

$$Si_2H_6\,(g) + h\nu \xrightarrow{k_4} H_3SiSi^*\,(g) + H_2\,(g) \qquad (6)$$

$$Si_2H_6\,(g) + h\nu \xrightarrow{k_5} H_3SiSiH_2^*\,(g) + H^*\,(g) \qquad (7)$$

$$SiH_2^*(g) + Si_2H_6\,(g) \xrightarrow{k_6} Si_3H_8\,(g) \qquad (8)$$

$$2\,SiH_3^*(g) \xrightarrow{k_7} Si_2H_6\,(g) \qquad (9)$$

Equations (3) through (7) occured through photolytic mechanisms. The quantum yields for the main Si fragments in reactions (3) through (7) are, 0.1, 0.4, 0.05, 0.1, and 0.05, respectively. This experimental information calls attention to the H_3SiSiH^* radical as the predominant precursor. This radical is also a relatively stable chemical species that reacts very slowly with silane and disilane. The radical H_3SiSiH^* suffers a 1,2 H isomerization to yield H_2SiSiH_2 that is the species transported to the substrate surface. Equations (8) and (9) were included to consider mechanisms for SiH_2^* and SiH_3^* disappearance. SiH_3^* is not expected to play a major role due to the fact that its quantum yield is too low compared to the H_3SiSiH^* radical.

Mass transport limitations are considered by using the corresponding mass balances. We use the fact that the laser beam produces only a minimal effect on the gas phase temperature [4].

Under the conditions of interest, relative "high pressures" (1 torr), the continuum approximation is still valid. For a Newtonian fluid, the velocity field is obtained by using the boundary layer approximation. The flow consists of an infinite stream flowing past a thin film of length L. The plate is consider infinitely wide, or z-axis symmetrical, and the flow is uniform at a horizontal approximation velocity, u_∞. The velocity vector has vertical and horizontal components inside the boundary layer only. An approximate solution for the flow problem has been presented by Whitaker [5]. The velocity in the horizontal direction (x-axis) is given by:

$$v_x = u_\infty \{\frac{3}{2}\frac{y}{\delta} - \frac{1}{2}(\frac{y}{\delta})^3\}, \tag{10}$$

Where δ is the boundary layer thickness and y is the vertical coordinate. The vertical velocity at the boundary layer is given by :

$$v_y = \frac{3}{8} u_\infty \frac{d\delta}{dx}, \tag{11}$$

and the boundary layer thickness can be calculated using,

$$\delta = 4.64\sqrt{\frac{x\,\nu}{u_\infty}}, \tag{12}$$

The gas temperature distribution is also calculated using boundary layer approximation, Whitaker [5],

$$T = T_W - (T_W - T_\infty)\{\frac{3}{2}\frac{y}{\delta_T} - \frac{1}{2}(\frac{y}{\delta_T})^3\} \tag{13}$$

T_∞ is the temperature outside the boundary layer, T_W is the temperature on the wafer wail, and δ_T is the thermal boundary layer thickness calculated from the momentum boundary layer thickness,

$$\delta_T = \delta\ 0.976\ Pr^{-0.333} \tag{14}$$

where Pr is the Prandt number for the gas phase ($c_p \mu/k$).

Equation (14) has been derived for the case where $\delta > \delta_T$. In our case $\delta_T >> \delta$. Calculation of the relation between both boundary layers thicknesses following the procedure reported by Coulson and Richardson [6] leads to similar results to equation (14).

Due to the relative high pressures considered in this work we assumed that the neutral species dominate the deposition rate. Mass balance equations are written for the species, Si_2H_6,H_2SiSiH_2, $SiH_2^*$and SiH_3^*, we do not take into account the radicals $Si(H_2)Si$ and Si_2H_5. These assumptions lead to the following mass balance equation,

$$(\underline{v} \circ \underline{\nabla}\, n_i) = \underline{\nabla} \circ \underline{\nabla}(D_{i,Si2H6}\, \underline{\nabla}\, n_i) + G_i = 0 \tag{15}$$

here n_i is the number of molecules of species i, $D_{i,Si2H6}$, is the binary diffusivity of species i in disilane. The net generation terms for gas phase reactions are derived using the method developed by Fogler [7], taking into account reactions (3) to (9) . The number of moles of active Si_2H_6 is computed from equation (2). Heterogeneous (surface) reactions enter in the material balance through the boundary conditions. We assume that the surface reactions for all the neutral species follow a Langmuir-Hinselwood mechanism (Fogler [7]),

$$H_2SiSiH_2\,(g) + 2\,S \overset{K_1}{\Leftrightarrow} S\text{-}H_2SiSiH_2\text{-}S \tag{16}$$

$$S\text{-}H_2SiSiH_2\text{-}S \xrightarrow{k_{s1}} 2\,Si\text{-}S + H_2\,(g) \tag{17}$$

The controlling step is given in each case by the hydrogen resorption, equation (17) . Then, the boundary conditions for equations (15) are given by

<u>BC1</u>. at $y = \delta_M$, for all x,

$$[n_i] = O \tag{18}$$

where $[n_i]$ represents concentration of Si_2H_4, SiH_2^* and SiH_3^*, respectively.

<u>BC2</u>. at x = L, for ally,

$$\frac{\partial n_i}{\partial r} = 0, \tag{19}$$

BC3. at x = O, for all y,

$$n_i = O, \tag{20}$$

BC4. at y = O, for all x,

$$D_{ni,Si2H6} \frac{\partial n_i}{\partial y} = k_{si}\,\beta(i)\,n_i \tag{21},$$

where $\beta(i)$ is the sticking coefficient onto the substrate surface. The hydrogen resorption rate constant (k_{si}) is calculated using the formula presented by Matsui et al. [8],

$$k_{si} = \beta(i) \sqrt{\frac{10^7\ R\ T}{2\ \pi\ M_i}}, \tag{22},$$

where R, M_i and T' are the gas constant, molecular weight, and temperature, respectively. This formula was derived assuming collisions between neutral- species and-the surface and involves the definition of a reactive sticking coefficient. It can be concluded from Buss et al. [9] work that both models are equivalent. The reactive sticking coefficient for radicals is considered to be one.

RESULTS AND DISCUSION

In order to implement the model described above a computer code was developed. Equations (15) describe a set of non-linear couple elliptical differential equations. An iterative over-relaxation finite difference scheme was used to solve equation (15) subject to the necessary boundary conditions. The numerical procedure calculates the concentration values using a five points scheme.

The velocity and temperature boundary layers are functions of the approximation velocity (v_∞) and the temperature outside the boundary layer (T_∞). Our calculations show that the temperature boundary layer is about one order of magnitude thicker than the velocity boundary layer. This fact is illustrated in Figure 2. Significant temperature effects on rate of thermal reactions can be expected because the gas temperature is significantly differently from the temperature outside the boundary layer at positions far away from the wafer. This result raises the issue of what temperature should be used to evaluate gas phase reactions. Tradittionall y in Chemical Vapor Deposition two different temperatures are referred to, the gas input temperature (T_∞) and the waffer temperature (T_w). The results shown in Figure 2 demonstrate the existance of a whole range of temperatures in between these two. This fact is not significant for relative low temperatures because under these conditions, photolytic and free radicals reactions determine the rate of deposition. The temperature effect becomes important for higher temperatures than the ones used in this work.

Solution of equation (15) allows us to determine concentration profiles for the different intermediates considered in this model. Figure (3) shows concentration profile for Si_2H_4 species. In this figure dimensionless concentrations are depicted. The disilane initial concentration was used to calculate the dimensionless concentrations in all cases. Figure (4) shows concentration profile for SiH_2^{**} species. These results show that the contribution of SiH_2^{**} to the total rate of deposition can not be neglected despite the fact that SiH_2^{**} concentrations are approximately four times lower than Si_2H_4 concentrations. This finding contradicts the assumptions made by Fowler et al. [4]. Similar results were calculated for several sets of different operating conditions.

The total rate of deposition was calculated by combining the individual rates of deposition of all the active species generated in the gas phase.

Figure (5) shows the rate of deposition profile calculated for the set of conditions used before. Significant non-uniformity can be appreciated. Results corresponding to direct disilane deposition rate mechanism are also presented. It can be concluded that for the temperature used in these calculation (523 K) this mechanism makes a negligible contribution to the total rate of deposition.

Comparisons of the results computed in this work and experimental data reported by Fowler et al. [4] showed good agreement between the computed and experimental results.

CONCLUSIONS

A model to compute silicon thin film deposition rates has been proposed. The model accounts for gas phase phenomena, convective transport effects and reaction rates on the solid wafer. One critical stage is the modeling of the velocity and temperature profiles in the proximity of the waffer. Boundary layer theory allowed us direct computation of velocity and temperature profiles. The velocity and temperature profiles were used to calculate the concentration profiles of all the important intermediate species. Mechanisms to explain the gas phase and solid surface reactions have been proposed. Despite their simplicity those mechanisms captured the main physical phenomena in LICVD. This research shows that significant temperature effects can be expected even at locations far away from the waffer. While Si_2H_4 makes the biggest contribution to the rate of deposition the contribution of other species can not be neglected. Good agreement between the model predictions and experimental results has been found.

Summarizing a model that can be used with certain confidence to predict Silicon deposition rates as a function of operating conditions has been reported.

REFERENCES

1). Sherman, A., *"heroical vapor deposition for microelectronics.* Noyes publications, Park Ridge, New Yersey, USA (1987).

2). Reinberg, A. R., *J. Electrochem. Sot. Extended Abstracts,* Vol. 1974-1, pg. 24.

3). Motooka, T. and Greene, J.E., *J.Appl. Phys., 59, 15(1986).*

4). Fowler, B.,Lian, S., Krishnan, S, Li, C., Samara, D., Manna, I. and Banerjee, S., *J. Appl. Phys.,* __73__, *8 12(1992).*

5). Whitaker, S., *Introduction to Fluid Mechanics,* R. E. Krieger Pub. Co., Melbourne, FL (1981).

6). Coulson, Richardson, *Chemical Engineering, Vol 1.,* Pergamon Press Pub. Co., London, New York, (1974).

7). Fogler, S., *Introduction to Chemical Reactor Design,* Prenttice Hail Pub. Co., New York, NY (1994).

8). Matsui, Y., Yuuki, A., Morita, N., and Tachibana, K., *Jpn. J. Appl. Phys.,* __9__, *1575(1987).*

9). Buss, R. J.; Ho, P.; Breiland, W. G. and Coltrin, M. E., "Reactive Sticking Coefficients for Silane and Disilane on Polycrstaline Silicon." *J. Appl. Phys.,* __63__, *2808(1988).*

Figures

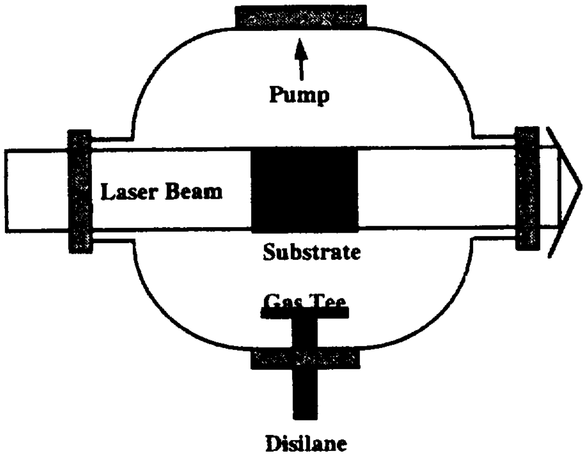


Fig.1. Photo-reactor for Si deposition from disilane

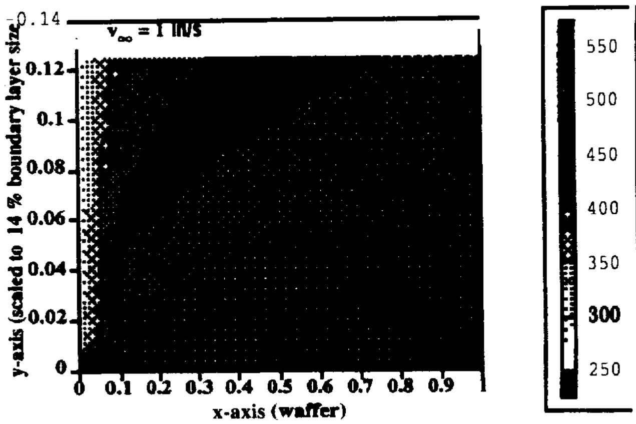


Fig. 2. Temperature contour.

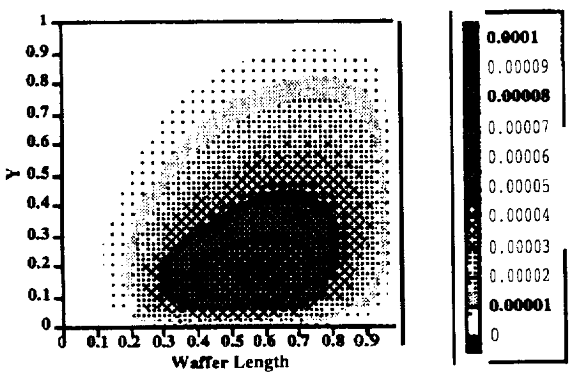


Fig. 3. Dimensionless Si_2H_4 concentration profile.

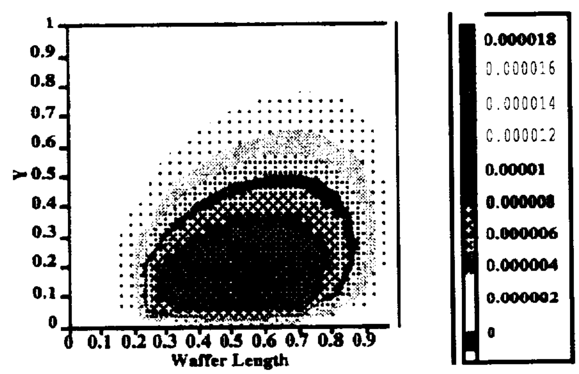


Fig. 4. Dimensionless SiH_2 concentration profile

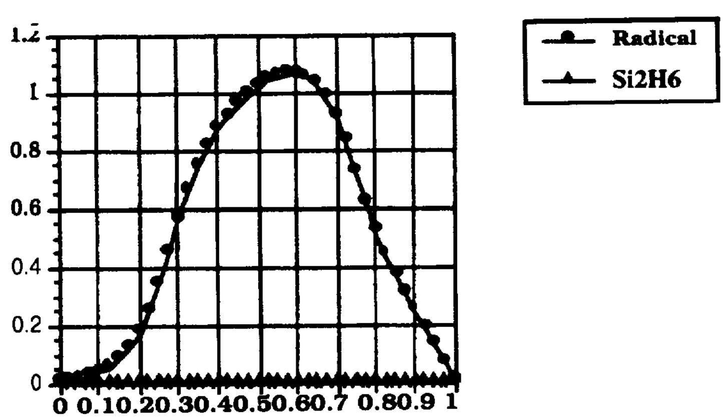


Fig. 5. Comparison of deposition rates caused by radicals and disilane (T= 523 K).

URC97139

SLIDING MODE CONTROL OF A SLEWING FLEXIBLE BEAM

N ASA

David G. Wilson[*] Gordon G. Parker[†] Gregory P. St arr[‡] Rush D. Robinett III[§]

2/33/63

Abstract

An output feedback sliding mode controller (SMC) is proposed to minimize the effects of vibrations of slewing flexible manipulators. A spline trajectory is used to generate ideal position and velocity commands. Constrained nonlinear optimization techniques are used to both calibrate nonlinear models and determine optimized gains to produce a rest-to-rest, residual vibration-free maneuver. Vibration-free maneuvers are import ant for current and future NASA space missions. This stud y required the development of the non-linear dynamic system equations of motion; robust control law design; numerical implementation; system identification; and verification using the Sandia National Laboratories flexible robot testbed. Results are shown for a slewing flexible beam.

Introduction

For NASA space applications, lightweight robotic manipulators are necessary to reduce launch costs, power consumption, and storage volume of the robot. Slewing structures with long flexible members, such as the Shuttle Remote Manipulator System (RMS) and the Space Station RMS, can excite vibrations. These vibrations can severely degrade the pointing accuracy, thus limiting the speed of rotation and productive use of current and future robotic systems. To achieve good control performance and position precision with current technology requires massive stiff manipulators. Since mass is the strongest driver of launch costs, massive telerobotic systems are unacceptable. Inherent flexibility for manipulator systems is a consequence of launch mass minimization. Flexibility is difficult to model but without its inclusion in the dynamic model, slewing performance will remain poor and cent rol marginally stable. The focus of this research is the development of a robust control system that demonstrates residual-vibration suppression and robust tracking using only colocated joint sensors and actuators. The methodology includes development of the dynamic system equations of motion; sliding mode control system design; optimized model matching and gain calculation; and experimental verification using the Sandia National Laboratories flexible robot testbed.

To find a practical feedback control for flexible arms, many researchers have investigated various control methods. A review of some of these techniques is given in Yeung and Chen.[1] Yeung and Chen also demonstrated successful feedback control of flexible arms using the sliding-mode technique. Nathan and Singh[2] developed a design approach for the control of a flexible robotic arm using variable structure system theory and pole assignment technique for stabilization. The closed-loop system was robust to variations in payload. Qian and Ma[3] have introduced variable structure sliding-mode technique for tip position control. The controller performance was evaluated through simulations. Choi and Shin[4] developed a sliding mode controller for tip position control of a single-flexible link manipulator subjected to parameter variation. Their algorithm showed fast and favorable system responses while maintaining low sensitivity y to imposed uncertainties. Parker and Robinett[5] developed an output feedback sliding mode control approach for nonlinear systems in general, with applications to flexible manipulators. Asymptotically stable sliding surfaces are specified in the output space. In addition, a constraint was derived, based on L yapunov's direct method, ensuring stability of the closed-loop system. The results presented in this paper is a realization of the output feedback sliding mode controller.

*Graduate Student, Department of Mechanical Engineering, The University of New Mexico, Albuquerque, NM 87131.
† professor, Michigan Technological Universit y, Houghton, MI 49931.
‡ Professor, Department of Mechanical Engineering, The University of New Mexico, Albuquerque, NM 87131.
§ Department Manager, Intelligent Systems Sensors and Control Dept., Sandia National Laboratories, Albuq., NM 87185.

Dynamic Model

The dynamic equations of motion for both the rigid body, θ and the flexible body, $q^i(t)$ degrees-of-freedom (DOF) are found using Lagrange's equations;

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right)-\frac{\partial L}{\partial \theta}=0 \tag{1}$$

$$\frac{d}{dt}\left(\frac{\partial L \partial L}{\partial \dot{q}^i}\right)-\frac{\partial L}{\partial q^i}=0 \tag{2}$$

The Lagrangian is, $L = T - V + W_F$, where the kinetic energy, T, the potential energy, V, and the work from external forces, W_F are defined as

$$T=\frac{1}{2}\bar{\rho}\int_0^L \dot{\vec{x}}^T \cdot \dot{\vec{x}}\, dx \tag{3}$$

$$V=\frac{1}{2}(EI)q^i q^j \int_0^L \phi^{i''}(x)\phi^{j''}(x)\, dx \tag{4}$$

$$W_F = \tau q^i \int_0^L \phi^i(x)\eta dx + \tau\theta \tag{5}$$

where i and j = 1 are the number of flexible DOF and x is the location along the beam. Applying Lagrange's equations results in the following nonlinear equations of motion;

$$\mathbf{M}(\mathbf{x})\ddot{\mathbf{x}} + \mathbf{N}(\mathbf{x},\dot{\mathbf{x}}) + \mathbf{K}(\mathbf{x},\dot{\mathbf{x}}) = \mathbf{B}(\mathbf{x})\mathbf{U} \tag{6}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} \tag{7}$$

where, x is an nx 1 vector of total DOF's, M is a nxn configuration dependant mass matrix; N is a $nx1$ vector of Coriolis and centripetal acceleration terms; K is a rum configuration dependant stiffness matrix including centrifugal stiffening terms; B is a nxm matrix of control weighting coefficients; U is a mx 1 vector of torque inputs; y is an $rx1$ vector of measurable outputs, C is an rxn matrix relating state variables to measurable outputs.

The dynamic equations of motion were developed using the method of quadratic modes.[6] Figure 1 shows a schematic of the slewing flexible beam defining the mathematical geometry.

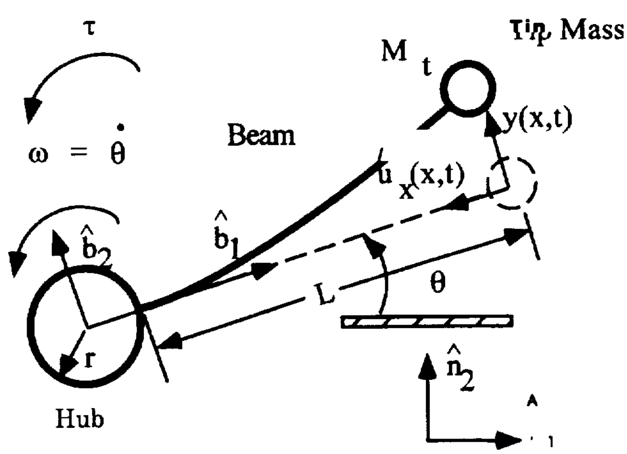


Figure 1: Slewing Flexible Beam Schematic

An expression for the deformation of a point along the beam is

$$\vec{u}(x,t) = u_x(x,t)\hat{b}_1 + y(x,t)\hat{b}_2 \tag{8}$$

where $\hat{b}_i$ are unit vectors associated with a moving coordinate system attached to the hub. The $\hat{n}_i$ unit vectors are associated with an inertial coordinate system. Define the following relationships for axial and transverse deflections as

$$u_x(x,t) = -\frac{1}{2}\int_0^x \left(\frac{d\,y(\xi,t)}{d\xi}\right)^2 d\xi \tag{9}$$

and

$$y(x,t) = \sum_{i=1}^{\infty} \phi^i(x) q^i(t)\,. \tag{10}$$

Equation (10) is assumed to be separable into $\phi^i(x)$, the mode shape basis functions and $q^i(t)$, the corresponding time-dependent generalized coordinates. The following equation gives the velocity of each point along the rotating beam's length:

$$\dot{\vec{x}}(x,t) = \frac{{}^N d}{dt}\left\{[r+x]\hat{b}_1 + \vec{u}(x,t)\right\}. \tag{11}$$

Performing the mathematical expansions and substituting the expressions for the kinetic energy, strain energy, and external work into Lagrange's equations (1,2), we arrive at the following equations for beam deflection and rotation, respectively.

$$\left\{EI\int_0^L \phi^{i''}(x)\phi^{j''}(x)\,dx - \left(\bar{\rho}\int_0^L \phi^i(x)\phi^j(x)\,dx + 2\bar{\rho}\int_0^L [r+x]g^{ij}(x)\,dx\right)\dot{\theta}^2\right\} q^i$$

$$+\left[\bar{\rho}\int_0^L \phi^i(x)\phi^j(x)\,dx\right]\ddot{q}^i + \left[\bar{\rho}\int_0^L [r+x]\phi^i(x)\,dx\right]\ddot{\theta} = \tau\int_0^L \phi^i(x)\eta dx \tag{12}$$

$$\left[\frac{1}{3}\rho L^3 + \rho r L^2 + \rho r^2 L + M_t(r+L)^2\right]\ddot{\theta} + \left[\bar{\rho}\int_0^L (r+x)\phi^i(x)\,dx\right]\ddot{q}^i = \tau \tag{13}$$

where $\bar{\rho} = \rho + M_t\delta(x-L)$ and ρ is mass per unit length.

Cantilever mode shapes given by Blevins[7], were used for this analysis and the quadratic modes for a beam were defined as:

$$g^{ij} = -\frac{1}{2}\int_0^x \phi^{i'}(\xi)\phi^{j'}(\xi)d\xi. \tag{14}$$

The final equations were arranged into the form of equations (6) and (7).

Sliding Mode Control

Sliding mode control provides an alternative to robot control with unknown parameters. The main advantage of SMC is its robustness to input disturbances once the sliding surface is reached. SMC uses a strategy whereby the active control law at any given time is chosen from a predefine set of control laws based on the current state of the system. SMC takes advantage of control law switching to move a system from an initial state to a prescribed surface in the state space. Once on that surface, a second control law is used to keep the state from leaving the surface while moving toward the desired final state. Using Lyapunov's direct method, SMC has been shown to be stable[5]. Furthermore, it is robust to model-parameter uncertainty and disturbances if bounds are known a priori. Sliding surfaces are identified for each sensor leading to the output feedback control law. A thorough development of SMC including several practical examples can be found in Utkin.[8]

The sliding surface may be chosen as

$$s = \mathbf{W}(\mathbf{y}-\mathbf{y_r}) + (\dot{\mathbf{y}}-\dot{\mathbf{y}}_\mathbf{r}) = 0 \tag{15}$$

where $\mathbf{y_r}$ is the desired sensor output time history and W is a positive definite matrix with real valued elements. The equivalent control is found by enforcing a condition of stationarity on the sliding surface,

$$\dot{s} = \mathbf{W}(\dot{\mathbf{y}} - \dot{\mathbf{y}}_r) + (\ddot{\mathbf{y}} - \ddot{\mathbf{y}}_r) = 0. \tag{16}$$

substituting for y from the equations of motion equations (6) and (7), intos and solving for $CM(x)^{-1}\mathbf{B}(\mathbf{x})\mathbf{U}$ yields;

$$\mathbf{CM(x)}^{-1}\mathbf{B(x)U} = \mathbf{CM(x)}^{-1}\mathbf{N}(\mathbf{x},\dot{\mathbf{x}}) + \mathbf{CM(x)}^{-1}\mathbf{K}(\mathbf{x},\dot{\mathbf{x}}) + \ddot{\mathbf{y}}_r - \mathbf{W}(\dot{\mathbf{y}} - \dot{\mathbf{y}}_r) \tag{17}$$

All terms involving x will be approximated with $\hat{\mathbf{x}}$, such that

$$\hat{\mathbf{x}} = \mathbf{C}^*\mathbf{y} \tag{18}$$

where C* takes on a form of a psuedo-inverse. In addition to substituting for the estimate of x the term $Atanh^{-1}(\beta s)$ is added to drive the output to the stable sliding surface ofs and the hyperbolic arctan is used to eliminate chatter through a boundary layer whose slope can be adjusted with /3. This results in the final output feedback sliding mode controller;

$$u = [\mathbf{CM}(\hat{\mathbf{x}})^{-1}\mathbf{B}(\hat{\mathbf{x}})]^{-1}[\mathbf{CM}(\hat{\mathbf{x}})^{-1}\mathbf{N}(\hat{\mathbf{x}},\dot{\hat{\mathbf{x}}}) + \mathbf{CM}(\hat{\mathbf{x}})^{-1}\mathbf{K}(\hat{\mathbf{x}},\dot{\hat{\mathbf{x}}}) + \ddot{\mathbf{y}}_r - \mathbf{W}(\dot{\mathbf{y}} - \dot{\mathbf{y}}_r) - Atanh^{-1}(\beta s)]. \tag{19}$$

Inversion is ensured by setting m = r. Stability has been established by using Lyapunov's direct method.[5] During the actual implementation of the SMC algorithm the following term was set to zero;

$$\mathbf{CM}(\hat{\mathbf{x}})^{-1}\mathbf{N}(\hat{\mathbf{x}},\dot{\hat{\mathbf{x}}}) + \mathbf{CM}(\hat{\mathbf{x}})^{-1}\mathbf{K}(\hat{\mathbf{x}},\dot{\hat{\mathbf{x}}}) = 0. \tag{20}$$

Optimization

A constrained optimization problem was formulated for the slewing flexible beam involving the physical parameters of the previously derived model and an experimental response to the trajectory input. Solving the trajectory optimization problem involved the use of a recursive quadratic programming alogrithm implemented in the MATLAB optimization toolbox.[9] A cost function of the form

$$J = W_1 \int_{t_0}^{t_f} \mathrm{e}_1^T \cdot \mathrm{e}_1 dt + W_2 \int_{t_0}^{t_f} \mathrm{e}_2^T \cdot \mathrm{e}_2 dt \tag{21}$$

subject to a number of inequality constraints $G(x) \leq 0$ was used for both the model matching and the optimized gain analysis.

The model matching optimization errors were set up as $\mathrm{e}_1 = (\theta_{model} - \theta_{exp})$ and $\mathrm{e}_2 = (\epsilon_{root_{model}} - \epsilon_{root_{exp}})$, where ϵ_{root} is the strain at the root of the beam. The first step involved setting the parameters belonging to the mass and stiffness properties of the hub and flexible link, The weights were set to $WI = 1.0$ and $W_2 = 0.0$. The optimizer was allowed to formulate error predominantly during the rise time $t_0 = 0.0$ to t_f = 0.35 seconds. After sufficient iterations the parameters would converge to nominal values. The second step was then to concentrate on the friction coefficients as parameters. The optimizer was set-up to work over the settling time portion of the trajectory from $t_0 = 0.35$ to $t_f = 0.6$ seconds until the parameters would converge to a nominal value. The third step used a representative set of parameters from both steps one and two. This set of parameters was allowed to only vary between ±20 percent. The optimizer worked over the combined time range of the trajectory from $t_0 = 0.0$ to t_f = 0,6 seconds. Upon successful convergence of these parameters, the rigid body portion of the single flexible link system is identified. The final step included setting $W_2 \geq 1.0$ and investigating parameters directly associated with the strain location and beam coefficients. This resulted in closer agreement to the experimental setup but was considered only a second-order effect.

For the controller optimization the errors are specified as $\mathrm{e}_1 = (\theta_c - 6)$ and $\mathrm{e}_2 = (0.0 - \epsilon_{root})$ where both variables # and ϵ_{root} are from the simulation model. The optimizer was set up for the hub angle from $to = 0.0$ to $t_f = 0.6$ seconds and for the root strain from $t_0 = 0.4$ to $t_f = 0.6$ seconds. The cost associated with the root strain is for after the maneuver is completed to minimize residual vibration. For all runs $WI = 1.0$ and $W_2 = 10$. To start out, large steps were taken to identify possible minimums. Starting with these minimums the step size was reduced until convergence. These gains were then implemented on the hardware to obtain experiment al responses,

Experimental and Numerical Results

The Sandia National Laboratories flexible robot testbed consists of modular flexible link/motor/hub mounting assemblies; elect ric DC motors and amplifiers; incremental encoders; bending strain gauges; and a dSPACE[10] real-time control computer and data acquisition system. The slewing flexible beam parameters are given in Table 1.

Parameter	Symbol	Value	Unit
Length	L	48.42	*cm*
Width	w	7.62	cm
Thickness	t	0.1574	cm.
Hub Radius	r	8.89	*cm*
Mass Density	ρ_m	2700	kg/m^3
Tip Mass	M_t	0.0	kg
Beam Stiff	EI	0.176	$kg.m^2$
Motor Inertia	J_m	6.92^{3}	$kg.m^2$
Viscous Damp	b_{vf}	1.37^{4}	$kg.m^2/s$

Table 1: Slewing Flexible Beam Physical Parameters

A numerical simulation was developed that realized the mathematical models developed earlier. MATLAB was used to implement the differential equations. The dynamics of the plant were treated as continuous states, while the control laws were treated as discrete states. All sampling was performed at 1000 Hz.

The reference motion trajectory of the hub is generated from a spline fit of the initial hub angle, equal to –90° to the final hub angle, equal to 90°, for the single flexible link case. The time for all the trajectory runs was specified as AT = 0.35 seconds.

Model Calibration/Matching

The goal of this section was to identify a model that best captured the dynamics of the actual system. A simple PD controller was used to slew the beam. By following the steps outlined in an earlier section, the following plots show the match between the model and the experimental set-up. Using empirically determined gains Figures 2 and 3 show the calibration plots for hub angle, hub velocity, root strain and mid-span strain, respectively.

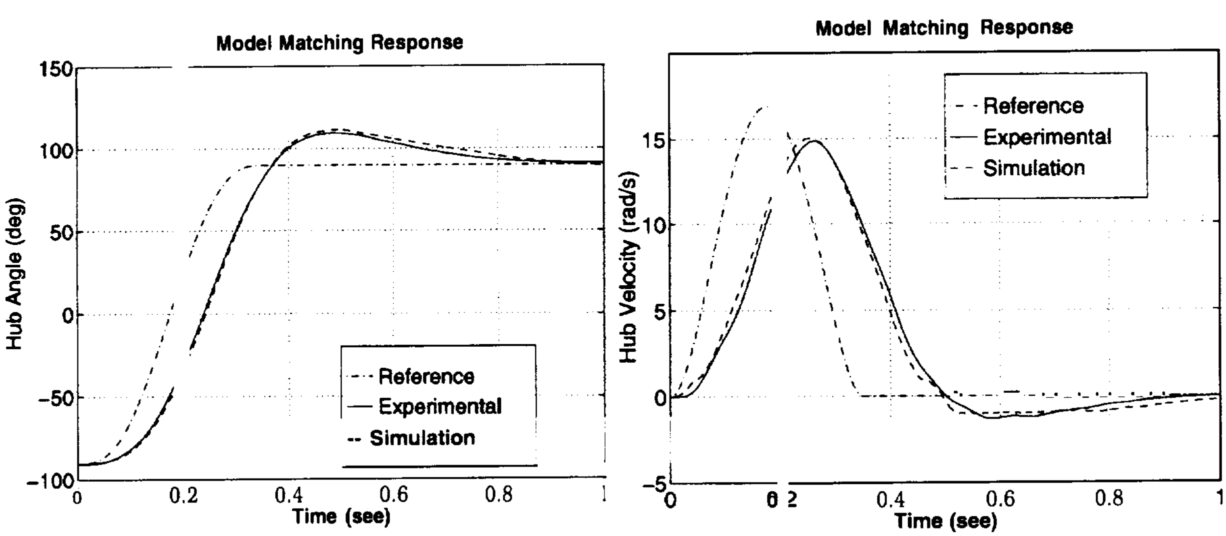


Figure 2: Hub Angle and Velocity Calibration Results

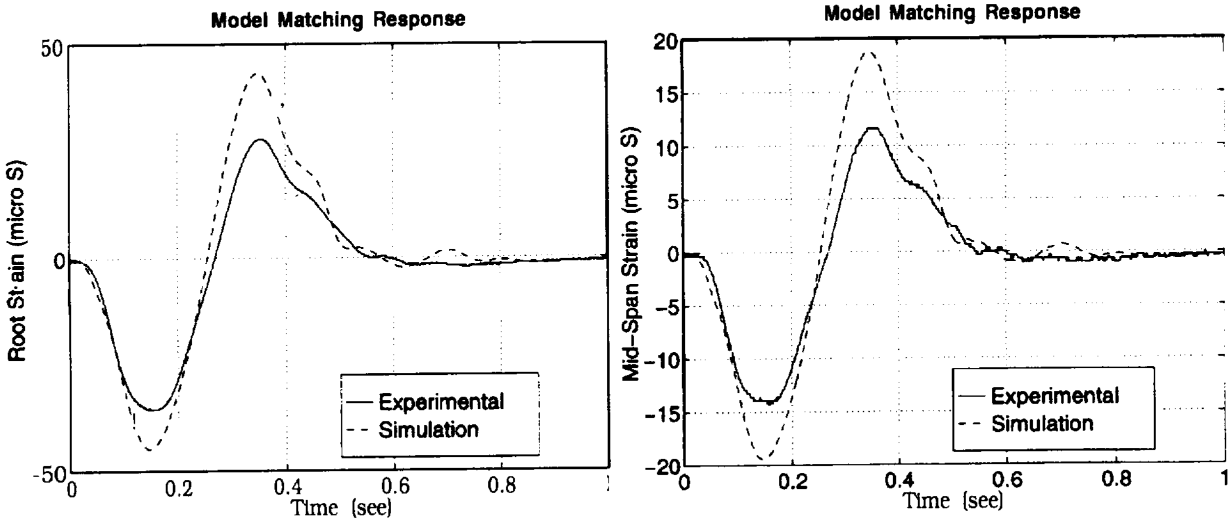


Figure 3: Root and Mid-Span Strain Calibration Results

Optimized Gains for Controllers

The calibrated model was used to predict the performance of the experimental set-up by using the gains determined from the constrained nonlinear optimization design. The results for the sliding mode control, where the W and A gains were optimized are shown in Figures 4 and 5, for the hub angle, hub velocity, root strain and mid-span strain, respectively.

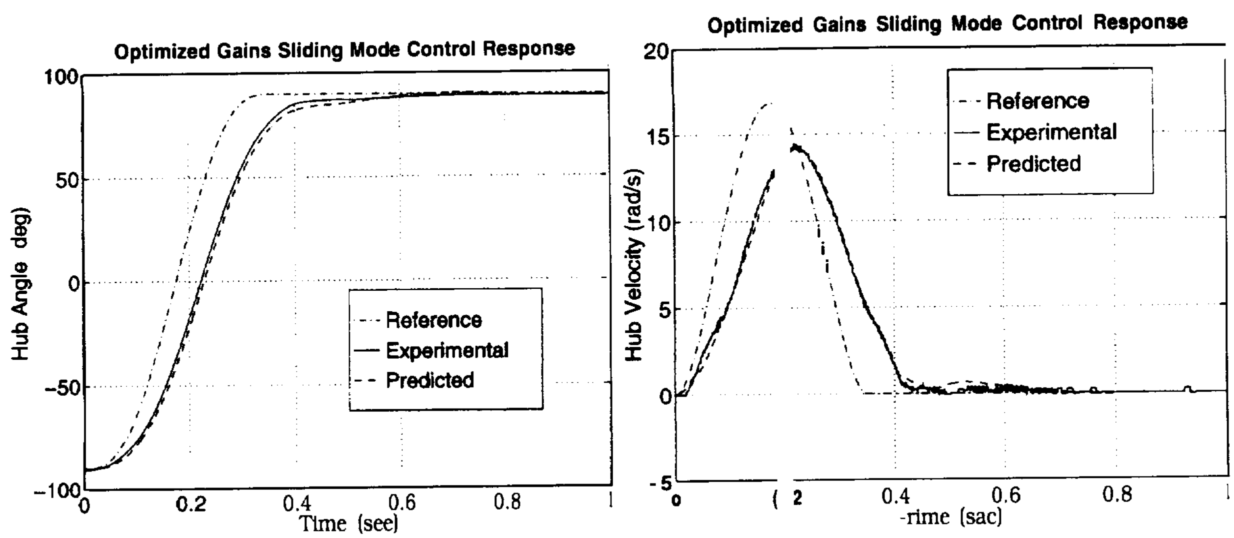


Figure 4: Hub Angle and Velocity Optimized SMC Results

Conclusions

A sliding mode controller was successfully demonstrated to minimize the effects of vibrations of slewing flexible beams. Optimization techniques were successful employed to determine meaningful nonlinear time domain models and optimized gain determination. In turn the optimized gains were used to predict flexible beam performance during large angle slews. These optimized gains were experimentally verifed on the Sandia National Laboratories flexible robot testbed. The SMC architecture showed minimum residual-vibration suppression and robust tracking using only colocated joint sensors and actuators. Future work will involve the use of 1) piezoceramic strain sensors and actuators[11,12] to enhance stability and tracking performance, and 2) the use of a two DOF planar flexible manipulator.

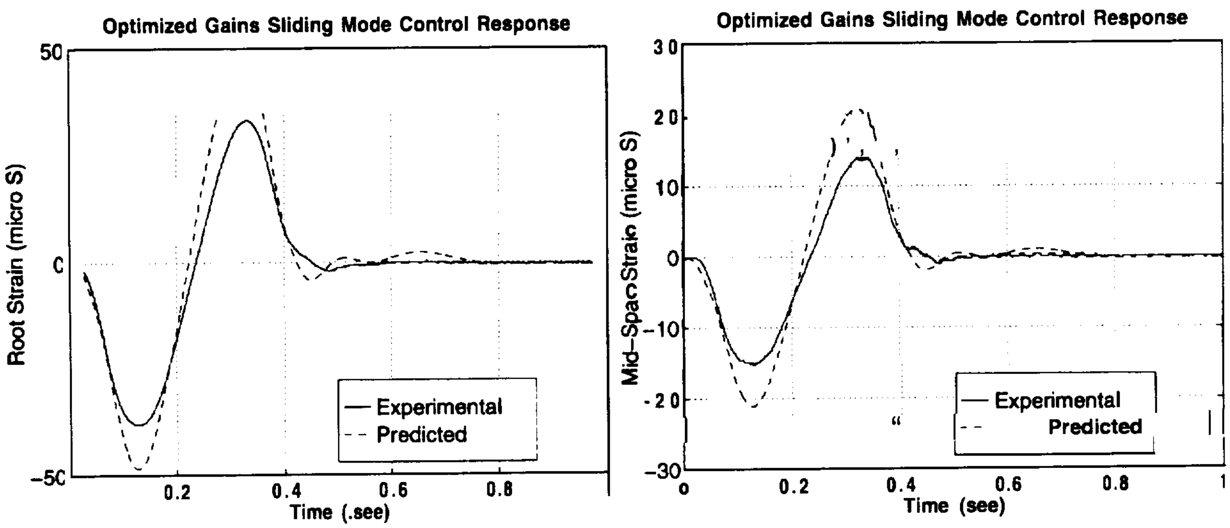


Figure 5: Root and Mid-Span Strain Optimized SMC Results

Acknowledgements

This work was supported by the U.S. Department of Energy under Contract DE-AC0494AL85000. The first author was also supported through a NASA ACE Fellowship.

References

[1]Yeung, K.S. and Chen, Y. P., "Regulation of a One-Link Flexible Robot Arm Using Sliding Mode Technique," *International J. of Control*,**49**, pg. 1965-1978, 1989.

[2]Nathan, P.J. and Singh. S. N., "Sliding Mode Control and Elastic Mode Stabilization of a Robotic Arm with Flexible Links," *J. of Dynamic Systems, Measurement and Control, 113,* pg. 667-676, 1991.

[3]Qian, W.T. and Ma, C. C. H., "A New Controller Design for Flexible One-Link Manipulator," *Transactions on Automatic Control, 37,* pg. 132-137, 1992.

[4]Choi, S.-B., Cheong, C.-C, Shin, H.-C., "Sliding Mode Control of Vibration in a Single-Link Flexible Arm with Parameter Variations", *J. of* Sound *and Vibration, 179(5),* pg 737-748, 1995.

[5]Parker, G. G., Robinett, R. D., "Output Feedback Sliding Mode Control with Application to Flexible Multibody Systems", Submitted to the *Journal of Robotic Systems, 1996.*

[6]Segalman, D. J., Dohrmann, C. R., "Dynamics of Rotating Flexible Structures by a Method of Quadratic Modes," Sandia National Laboratories, SAND90-2737, December 1990.

[7]Blevins, R.D., *Formulas for Natural Frequency and Mode* Shape, Krieger Publishing Company, 1993.

[8]Utkin, V. I., *Sliding Modes in Control Optimization,* Springer-Verlag, 1981.

'Grace, A., *Optimization TOOLBOX for Use with MA TLAB,* The MathWorks, Inc., 1995.

[10]dSPACE, Real-Time Control System, dSPACE digital signal processing and control engineering GmbH, Technologiepark 25, D-33100 Paderborn, Germany.

[11]Warren, S. R., Voulgaris, P.G, Bergman, L. A., "Robust Control of a Slewing Beam System", *J. of Vibration and Control,* Vol 1: 251-271, 1995.

[12] Wilson, D. G., Searle, I., Ikegami, I., and Starr, G. P., " Dynamic Characterization of Smart Structures for Active Vibration Control Applications," *ASME Winter Annual Meeting, Symposium on Vibro-Acoustic* Applications, Nov. 17-22, 1996, Atlanta, Ga.

MONITORING LAND SURFACE SOIL MOISTURE FROM SPACE WITH IN-SITU SENSORS VALIDATION - THE HUNTSVILLE EXAMPLE

Steve Shih-Tseng Wu

NASA/MSFC/ES4 1 Global Hydrology and Climate Center
977 Explorer Boulevard, Huntsville, AL 35806
T: 205.922.5961, F: 205.922.5723, e-mail: steve.wu@msfc.nasa.gov

ABSTRACT

Based on recent advances in microwave remote sensing of soil moisture and in pursuit of research interests in areas of hydrology, soil climatology, and remote sensing, the Center for Hydrology, Soil Climatology, and Remote Sensing (HSCaRS) conducted the Huntsville '96 field experiment in Huntsville, Alabama from July 1-14, 1996. We, researchers at the Global Hydrology and Climate Center's MSFC/ES41, are interested in using ground-based microwave sensors, to simulate land surface brightness signatures of those space borne sensors that were in operation or to be launched in the near future. The analyses of data collected by the Advanced Microwave Precipitation Radiometer (AMPR) and thc C-band radiometer, which together contained five frequencies (6.925, 10.7, 19.35, 37.1, and 85.5 GHz), and with concurrent in-situ collection of surface cover conditions (surface temperature, surface roughness, vegetation, and surface topology) and soil moisture content, would result in a better understanding of the data acquired over land surfaces by the Special Sensor Microwave Imager (SSM/I), the Tropical Rainfall Measuring Mission Microwave Imager (TMI), and the Advanced Microwave Scanning Radiometer (AMSR), because these spaceborne sensors contained these five frequencies. This paper described the approach taken and the specific objective to be accomplished in the Huntsville '97 field experiment.

Key **Words:** Soil moisture, Microwave remote sensing, In-situ sensors, Spaceborne imagers.

INTRODUCTION

Advancements of microwave remote sensing technology, such as a suite of space-borne active and passive microwave sensors become operational or to be operational in next few years, makes monitoring land surface soil moisture a possibility. On the other hand, recent advances in microwave researches indicated that both passive microwave and active microwave techniques have provided solid theoretical and experimental results that the top five cm of soil moisture can be measured from ground-based truck, aircraft and space platforms under a variety of environmental conditions and through a moderate vegetation cover (Engman. 1995). We, researchers at the Global Hydrology and Climate Center, are interested in the microwave sensors with frequencies corresponding to those current or near future space-borne active and passive microwave sensors (Wu, 1996). We also address several of the specific recommendations of the workshop attendees of a NASA's Office of Mission to Planet Earth (MTPE) sponsored workshop in 1994 on soil moisture (Wei, 1994). It is also consistent with the overall goals of MTPE Strategic Enterprise as it pertains to studies involving land-cover change, global productivity, and long-term climate variability.

At the same time, We are one of the research teams participated in the Huntsville '96 field experiment in microwave remote sensing of soil moisture in Huntsville, Alabama from July 1-14, 1996, sponsored by the Center for Hydrology, Soil Climatology, and Remote Sensing (HSCaRS). The remote sensing measurements were supported by soil profile instrument systems, gravimetric moisture measurements, and soil and vegetation characterization. Additionally, radiation, wind, air temperature, and relative humidity measurements were also included, Scientific objectives focused on defining the soil depth emitting and reflecting energy at various microwave wavelengths; characterizing temporal and spatial variability of surface moisture, and studying the capability of measuring moisture at different frequencies. Both Huntsville '96 and Huntsville '97 field experiments focus on a small-scale (plot-size) testbed with well equipped *in-situ* instruments and three microwave soil moisture remote sensing systems. Preliminary results of the Huntsville '96 field experiment will be presented in the American Meteorological Society's Annual Meeting on February 2-7, 1997 (Laymen et al., 1997). This paper, with encouraging findings of Huntsville '96 as the starting point, will address several issues for the Huntsville ' 97 field experiment:

1. We will conduct the well controlled concurrent active/passive microwave data collections, with similar frequencies of L-, C-, and X-band, over the testbed for a range of conditions.
2. To facilitate and to convert the Advanced Microwave Precipitation Radiometer (AMPR) to become suitable for ground-based applications, we will modify the AMPR system's hot and cold loads calibration subsystem and the data acquisition subsystem.

3. Since the Special Sensor Microwave Imager (SSM/I) has been in operation nearly 10 years with its three (19.35, 37.1, 85.5 GHz) frequency-channels have been duplicated by the AMPR and the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI) have the 10.7 GHz frequency-channel and to be launched in 1997, we will collect the AMPR data, with site validation objective, to simulate the SSM/I and TMI sensors' land surface brightness temperatures, T_B, to get a better understanding and interpretation of these sensors' acquired data over land surfaces.
4. Since the Advanced Microwave Scanning Radiometer (AMSR) to be launched on the EOS PM-1 platform in the early 2000s have the lowest frequencies of 6.925 GHz, we will also conduct data collection using the 4-8 GHz step frequency C-band radiometer in the Huntsville '97 field experiment to simulate AMSR land surfaces brightness signatures.
5. Since the ground-based AMPR system is able to measure a half space, i.e. from nadir (looking directly downward from the boon) to zenith (looking directly upward) with 360 degree horizontal rotation, the measured sky and land surfaces brightness temperatures will be used to simulate the space borne microwave sensors' microwave signatures over land, the TRMM microwave imager in particular. During thunder storms or rain conditions, the upward looking AMPR will be able to collect rain rate over land just as the rain gage.
6. We will start developing the AMPR and C-band collected brightness temperature data bases for which the space borne microwave sensors' measurements over land surfaces would be validated.

TESTBED AND IN-SITU INSTRUMENTS

A research testbed was established at Alabama A&M University's Agricultural Research Station located about 20 km north of Huntsville, Alabama (Fig. 1). The testbed consisted of four plots about 50 x 60 m. Two plots were bare of vegetation and two others were vegetated. One of the vegetated plots had a tall fescue cover, whereas the other had a mixture of vegetation. A total of 110 soil cores (1 m) were extracted from the testbed on a 10 m grid for soil characterization. Soil of the grass-covered plot is classified as clay loam to silty clay loam, whereas the other plots are silt loam. Clay content increases with depth to 1 m in all plots from 24% to about 50'%.. Clay content is slightly higher in the grass plot. The organic matter content in the surfacc 15 cm is less than 2%. The configuration of the testbed will be the same in 97 experiment with cotton to replace the rough bare field.

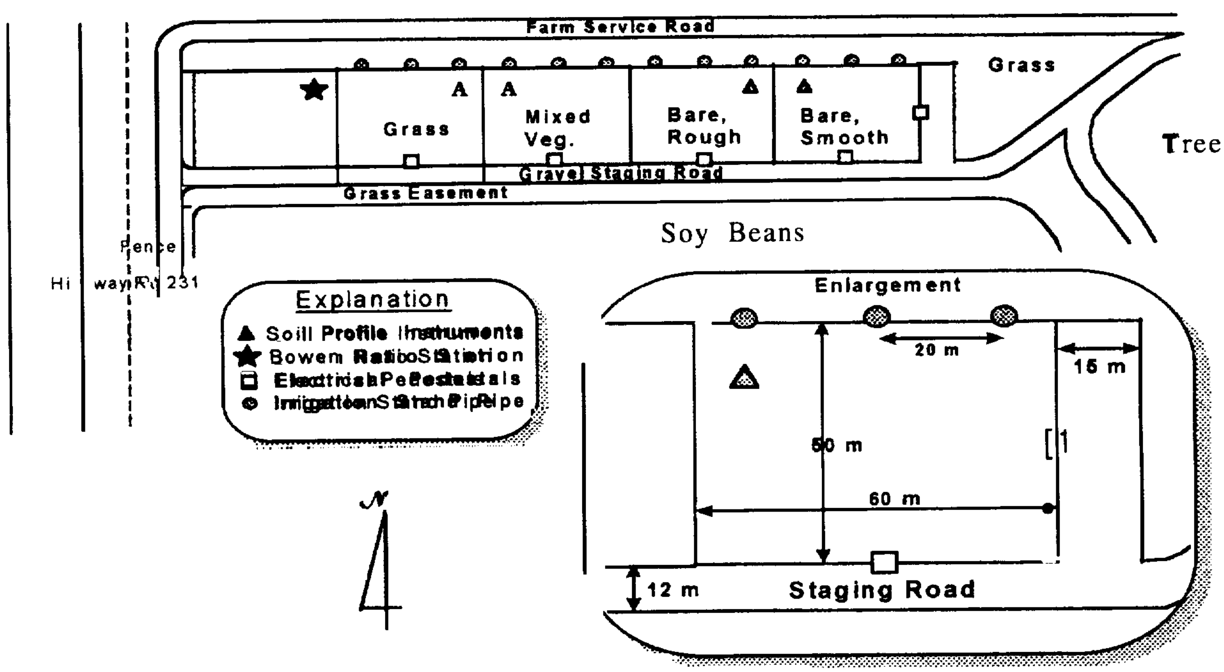


Figure 1. Diagram showing the layout of the testbed. Remote sensing instruments staged from the gravel road on the south side of the plots.

In the Huntsville '96 field experiment (Laymon et al., 1997):

Soil samples for gravimetric moisture content were collected from the four experimental plots at random locations and occasionally on a 10 x 10 m grid. The soil was sampled at five depth intervals (0-1, 0-3, 0-5, 0-7, and 0-10 cm) in the morning (~08:00) and at a single depth interval (0-5 cm) in the afternoon (~14:30). Standard

procedures were used to analyze these samples. Soil bulk density of the upper 15 cm was determined using the excavation method developed by the U.S. Department of Agriculture. Bulk density samples were collected before irrigation (DOY183), two days after irrigation (DOY186, 187), and after a one week dry-down period (DOY194). Figure 2 shows the daily amount of precipitation (irrigation, rainfall) measured by distributed rain gages. Irrigation totaling 34.3 mm was applied on DOY184- 186. Two rain events occurred totaling 43.7 mm; 36.1 mm fell on DOY189-190 and 7.6 mm fell on DOY196 (Fig. 2). Irrigation was applied to bare and vegetated plots at slightly different times. Because gravimetric sampling is destructive, samples were only acquired from the northern half of the plots outside the area sampled by the remote sensing instruments. To evaluate potential errors resulting from sampling different halves of the plots, we conducted an assessment of the spatial variability of surface moisture for each plot. Results show that, in general, the variance in gravimetric moisture content was low. Peaks in variance occurred on two days for the rough bare plot and on three days for the mixed vegetation plot. For the most part, soil moisture behaved as expected in response to the two dominant wetting events (Fig. 3).

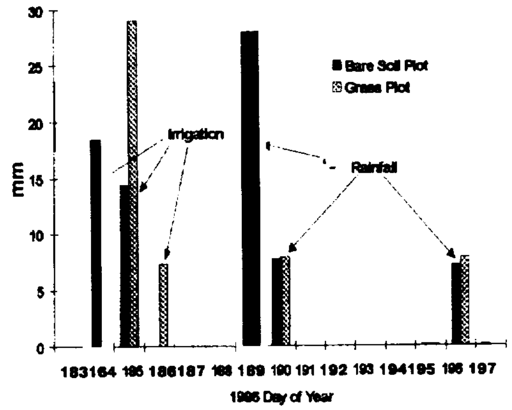


Figure 2. The daily amount of precipitation (irrigation, rainfall) measured by distributed rain gages.

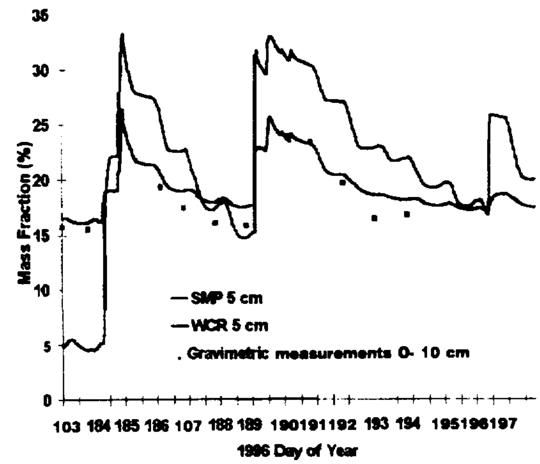


Figure 3. The soil water content.

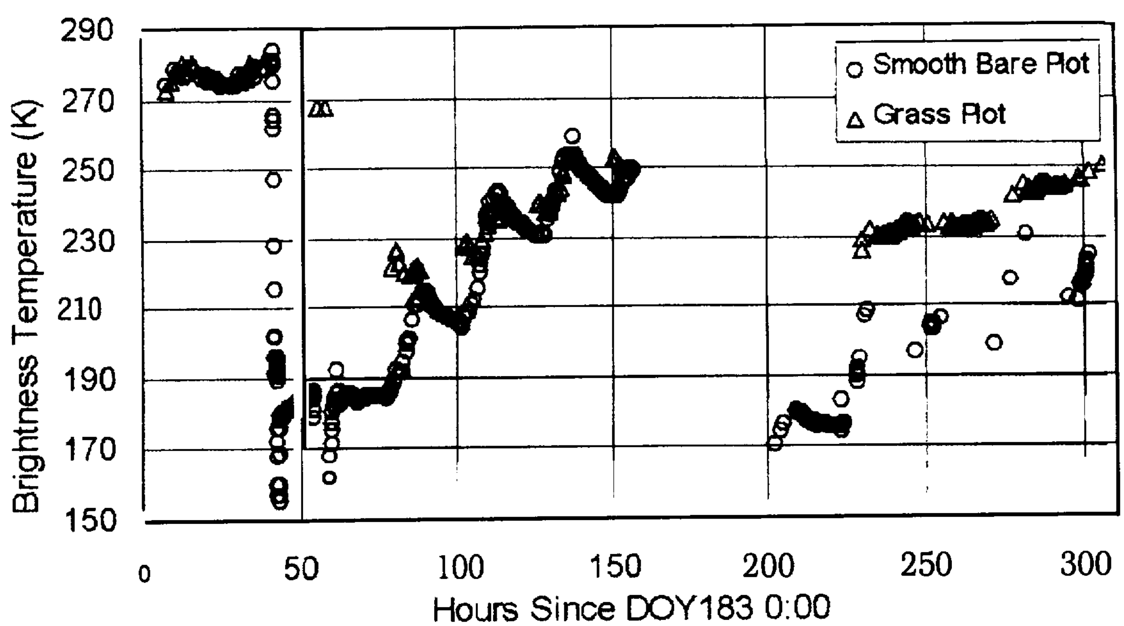


Figure 4. Time series of L-band microwave brightness temperature for the smooth bare and grass plots.

Surface soil moisture was about 7% as the experiment began. It rose rapidly to about 32% after the wetting events and dried slowly afterwards. Some differences in the moisture response among plots were observed. In order to obtain information required for application and testing of radiative transfer algorithms and land surface models, several vegetation properties were measured during the field experiment including vegetation height, wet and dry biomass, dielectric constant (for grass plot only), roughness of underlying soil, leaf dimension and orientation, and the angular distribution of stems and leaves. Except as noted, each variable was measured on both vegetated plots on several occasions during the two-week experiment period.

The microwave remote sensing data collected during this experiment lend themselves to an evaluation of the radiometer and radar performance in a) capturing the time-series of moisture change through two wetting and drying cycles, b) capturing the diurnal cycle of moisture change, c) comparison among various microwave wavelengths in a and b, and d) comparison of a through c as a function of different vegetation cover. At the outset of the experiment, microwave brightness temperatures (T_B) for the L and S bands were between 250° to 280° K for all plots (Fig. 4). After the initial wetting, T_B on the smooth bare plot dropped to 180° K. T_B on the vegetated plots only dropped by about 30° K (Fig. 4). T_B recovered to about 80% of its initial value in the four days prior to the second wetting on DOY 189. L-band T_B's are commonly higher than those of S-band. The magnitude of these differences varies for the different land cover conditions. The inter-band T_B's are most similar on the smooth bare plot and least similar on the grass plot.

In the Huntsville '97 field experiment:

The mode of the AMPR data collection and its close association with the C-band radiometer's data sets will be modified to better simulate the space borne microwave sensors' brightness signatures over land surfaces. The testbed and the *in-situ* instruments will stay the same as described in previous section; the field data collection scheme will also be the same.

TRUCK RADARS AND TRUCK RADIOMETERS

The Functionality and Operation ofMicrowave Soil Moisture Sensors

Three separate microwave remote sensing systems were deployed during the Huntsville '96 field experiment and also will be deployed next year, each with the capability to measure over multiple frequencies:

1. The S- (2.65 GHz) and L-band (1 .413 GHz) microwave radiometers (SLMR) with 15° beamwidth, 0.1° K radiometric resolution, and 1 sec. integration time were integrated into one system (Jackson et al., 1995). A step frequency (4 -8 GHz) C-band radiometer with six channels: 4.63, 5.06, 5.91,6.34, 6.77, 7.20 GHz 15° beamwidth, 0.1 K radiometric resolution, and one second integration time. The antennas of the SLMR were mounted to observe horizontal polarization. Data were acquired at a look angle of 15° from nadir and at a nominal height of 14 m. The radiometers sampled about 220 times each second. Each measurement was made using a 15 second integration time. An "autocollect" capability allowed us to leave the system unattended over a site for extended periods. This capability was utilized to collect data throughout the night, during irrigation, and at other times of interest. Measurements were made at each plot on nearly an hourly basis during the day throughout the experiment. The radiometers were placed in autocollect mode over the smooth bare plot at night during the first week and over the grass plot during the second week. Radiometers were calibrated over a pond at the research station before and after the experiment. An entire day was dedicated to cycling through absorber, sky and water measurements. Over water, measurements were made in 5° increments from 5° to 65°. Surface water temperature was measured throughout the day with three temperature sensors that were floating 2 cm below the pond surface.
2. The Advanced Microwave Precipitation Radiometer (AMPR) comprised another system (Spencer et al., 1994). AMPR is a 4-channel (10.7, 19.35, 37.1, 85.5 GHz) continuous scanning instrument previously deployed on the NASA ER-2 aircraft. The 3 dB beamwidths of the four channels are 8.0°, 8.0°, 4.2°, 1.8°, respectively. The reflector scans 90° while sampling 50 beam spots every 1.8° over 2.5 sees. After four scans, the reflector scans up to measure internal warm and cold calibration loads. The feed-horn is fixed with respect to the reflector, thus providing H and V polarization at opposite ends of the scan. Between the two scan extremes, the two polarization states are combined and are equal at the middle of the scan. The instrument was offset 45° from the other radiometers so that H polarization coincided with the look angle of the SLMR.
3. The third system was comprised of an L- (1.6 GHz), C- (4.75 GHz), and X-band (10.0 GHz) radar with 9°-120 beamwidth at 3 dB and 4 polarization combinations (O'Neill et al., 1995). The radars were deployed from another boom track and instrument control and data acquisition systems were operated

from the back of the truck. Radar data were acquired at two incidence angles as measured from nadir—at 15° for coincidence with the radio-meter measurements, and at 45° for vegetation modeling. The radar system used two modes of operation, "sweep" and "snapshot." In sweep mode, the radar boom was slowly rotated through 120° of azimuth in order to acquire a spatial average across the test plot. This was the standard mode during the day. In snapshot mode, the radar boom was stationary and a single radar footprint was imaged repeatedly over time. This mode was automated and used to collect data throughout the night. During the first week of the experiment, radar back-scatter measurements were made in sweep mode on all four test plots each day, with collection cycles initiated at 06:00, 10:00, and 14:00 hours local daylight time. At night, the radar truck was placed in autocollect snapshot mode over the Bare-smooth plot to acquire data coincident with the radiometers. After rain events, the radar truck staged at a single plot on a given day in order to get complete diurnal data for each plot. Hourly sweep data were collected between 08:00 and 16:00, and half-hourly snapshot data over night between 17:00 and 07:00.

The AMPR Modification for the Huntsville '97

To facilitate and to convert the AMPR system to become suitable for ground-based applications, it is necessary to modify its hot and cold loads calibration subsystem and the data acquisition subsystem

Hot and Cold Loads Calibration System: The AMPR is a total power radiometer. Because of this feature, a hot and cold loads are needed to conduct the system calibration by means of scanning through the hot and cold loads for every four scans of the one data acquisition cycle. Frequent comparison of the measured land surfaces brightness temperatures with respect to the hot and cold load temperatures requires constancy and a large spread of temperature between the hot and cold loads. The hot load was designed to heat up and keep it at a constant temperature of 320K. The cold load was originally designed to use the air temperature at 20km flight altitude of the ER-2. At that altitude the air flow through the cold load is usually well below freezing point of 210k. A 110K or more spread of temperature between the hot and cold loads is sufficient to do the calibration. When AMPR was ground based as it was used in the "Huntsville 96" field experiment, the above ground air temperature in July not only subject to day-night variation, the noon time temperature can exceed 25C (295 K). Without modification of AMPR cold load, the calibration system produce only 15K temperature spread instead of the 110K it needed. Therefore, we do need to modify the cold load by using a freezing system to cool it down to well below freezing point (21 OK) to make it comparable to the original design specification.

Data Acquisition System: The AMPR system was originally designed and manufactured by MSFC for airborne (ER-2), large area coverage, and very long automated data acquisition flight times, its data acquisition system with very burdensome flight data included is not suitable to be used in the ground-based AMPR system. Lesson learned from the "Huntsville 96" microwave soil moisture measurement field experiment indicates a completely different data acquisition system is needed for the ground-based AMPR system. The switch-on completely automated data acquisition scheme, used in the past or future AMPR ER-2 borne flight missions, was designed to meet the requirement of ER-2 platform, because there is only one man pilot. The pilot have no time to pay attention to his pay load instruments for which AMPR is just one part of the whole pay load. Therefore, all sensors on ER-2 platform must be fully automated; the pilot turn on the instruments and forget about it during the flight mission. For the ground based AMPR we will have an operator to operate the system and he shall decide how to conduct the measurement, e.g. measuring the soil moisture of a test plot every 15 minutes or every hour to see its change. We do not need to measure soil moisture or whatever parameters all the times. We will implement a controllable data acquisition scheme and make it operable on and off by the operator and at the same time keep the hot and cold load calibration system running all day long to maintain constancy of the Calibration temperature. This scheme can be made by using a separated power source for the calibration system. Computer programs will be developed and used to operate the data acquisition system and with operator interactive or interfacing capabilities.

THE AMPR AND C-BAND RADIOMETER'S ROLE IN SOIL MOISTURE SENSING

It is believed that by using the AMPR and the C-band radiometer will result in a better understanding of the data acquired over land surfaces by the Special Sensor Microwave Imager (SSM/I), the Tropical Rainfall Measuring Mission Microwave Imager (TMI), and the Advanced Microwave Scanning Radiometer (AMSR).

A modification of the AMPR subsystems as described in previous section by itself will have the following benefits:

1. The developed truck-mount ground-based AMPR system provides a better position to participate in the "Huntsville 97" field experiments, because ER-2 borne AMPR are very costly.

2. *The* three frequencies of Tom Jackson's microwave system, and four frequencies of the AMPR system to be used to collect microwave brightness temperature signatures (in 1.4, 2.65, 4 - 8, 10.7, 19.35, 37.1, and 85.5 GHz) over the Huntsville '97's selected test-plots would provide a better understanding of soil moisture over these test-plots.
3. On the other hand, since the Advanced Microwave Scanning Radiometer (AMSR) to be launched on the EOS PM-1 platform in 2000 have the lowest frequencies of 6.925 GHz, using the 4-8 GHz step frequency C-band radiometer to collect data in the Huntsville '97 field experiment will enable us to simulate AMSR's land surfaces brightness signatures.

CONCLUSION

In the Huntsville '96 field experiment, three frequencies of Tom Jackson's microwave system, and four frequencies of the AMPR system were used to collect microwave brightness temperature signatures (in 1.4, *2.65*, 4 -*7*, *10.7*, *19.35*, *37.1*, and *85.5* GHz) over three test-plots of bare soil, alfalfa, and grass, respectively. A quick look of the collected data indicated that significant decrease of brightness temperature was caused by the increase of soil moisture which is in agreement with the theory. A team work approach of the Huntsville '96 field experiment resulted in a valuable data sets and the follow on analyses would provide a better understanding of soil moisture over that three test-plots. Analyses of the AMPR collected data together with concurrent in-situ collection of surface cover conditions (surface temperature, surface roughness, vegetation, and surface topology) and soil moisture content, would result in a better understanding of the data acquired over land surfaces by the Special Sensor Microwave Imager (SSM/I), the Tropical Rainfall Measuring Mission Microwave Imager (TMI), and the Advanced Microwave Scanning Radiometer (AMSR), because the AMPR and the C-band radiometer contained four frequencies (6.925, 10.7, 19.35,37.1, and 85.5 GHz) of these spaceborne sensors.

ACKNOWLEDGMENTS

This work was supported Grant No. NCCW-0084 from the National Aeronautics and Space Administration (NASA), Washington, DC. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U. S. Government. Acknowledgment is given to the Center for Hydrology, Soil Climatology, and Remote Sensing (HSCaRS) support staff. Contribution from HSCaRS the Department of Plant, Soil and Animal Sciences and the Agricultural Experiment Station, Alabama A&M University, Normal, AL 35762. Journal No. 346.

Acknowledgment is also given to Drs. M-Y Wei, Jim Arnold, Roy Spenser, and Tom Jackson for the fruition of using the Advanced Microwave Precipitation Radiometer in Huntsville '96 experiment.

REFERENCES

Engman, E.T., 1995: Microwave remote sensing of soil moisture. *Proc. Int. Geosci. Rem. Sens. Symp., IEEE No. 95CH35770*, Florence, Italy, July 10-14, Vol. I, p. 489-491.

Jackson, T.J., P.E., O'Neill, W.P. Kustas, E. Bennett, C.T. Swift, 1995: Passive microwave observations of diurnal soil moisture at 1.4 and 2.65 GHz, *Proc. Int. Geosci. Rem Sens. Symp., IEEE No. 95CH35770*, Florence, Italy, July 10-14, Vol. I, p. 492-494.

Laymen, C., W. Belisle, T. Coleman, W. Crosson, A. Fahsi, R Hood, T. Jackson, J. Luvall, A. Manu, P, O'Neill, Z. Senwo, T. Tsegaye, S. Wu, 1996: Huntsville '96 microwave remote sensing of soil moisture experiment: overview and preliminary results. *Proceedings of the Conference on Hydrology, Annual* Meeting, American Meteorological Society February 2-7, 1997, Long Beach, CA

O'Neill, P.E., J.J. Petrella, and A.Y. Hsu, 1995: Comparison of mulitfrequency truck radar and SIR-C backscatter for soil moisture estimation in Washita '94. *Proc. IGARSS '95, IEEE*, Florence Italy, July 10-14, Vol. I, p. 368-370.

Spencer, R.W., R.E. Hood, F.J. LaFontaine, E.A. Smith, R. Platt, J. Galliano, V.L. Griffin, and E. Lob], 1994: High-resolution imaging of rain systems with the advanced microwave precipitation radiometer. Jour. Atmos. Oceanic Tech., 11, 849-857.

Wei, M_Y., 1994: "Soil moisture: report of a workshop in Tiburon, California 25-27 January 1994," NASA Conference Publication 3319, NASA Headquarters.

Wu, S. T., 1996: Microwave remote sensing of land surfaces soil moisture at Global Hydrology and Climate Center. *Proc. Int. Geosci. Rem. Sens. Symp., IEEE No. 96CH35875*, Lincoln, Nebraska, May 27-3 I, Vol. 2, pp. 1309-1311.

/// / 7/6

URC97141

Characteristics of Defects and Microstructure Of Tungsten Ton Implanted Pure Iron

Dehua Yang Jianren Zhou

Center for Applied Radiation Research/
Department of Mechanical Engineering
Prairie View A & M University
Prairie View, TX 77446

Abstract

Ion implantation, ion sputtering and radiation damage consist of the important aspects of ion-solid interactions. In ion implantation process, collision between energized ions and atoms at crystal lattice and rearrangement of atoms result in defect formation and phase transformation, which may cause significant changes of materials in both compositions and structures. These changes could modify and improve the various properties of materials. Therefore, the research in this area is important for both the theoretical understanding of ion-solid interactions and technology applications in aerospace andother industries. In present work, tungsten ion implantation into pure iron at dose of 3×10^{17} ions/cm^2 was performed using a metal vapor vacuum arc ion source implanter. The chemical composition and microstrucural changes caused by tungsten ion implantation were examined by means of Auger electron spectroscopy (AES), X-ray photoelectron spectroscopy (XPS) and transmission electron microscopy (TEM). The analytic test results indicated that tungsten ion implantation changed the microstructure of pure iron from single phase α-Fe into complicated microstructures composed of et-Fe solid solution of W, amorphous Fe-W-C and precipitate of Fe_3C. The formation of the amorphous phase was the results of lattice distortion and grain rupture due to tungsten ion collision and injection, The formation of carbide phases was related to carbon incorporation in implantation process.

1. Introduction

Ion implantation, ion sputtering and radiation damage consist of the important aspects of ion-solid interaction. The research in these areas are important both for the theoretical understanding of ion-solid interactions and for technology applications in aerospace and other industries. In ion implantation process, collision between energized ions and atoms at crystal lattice and rearrangement of atoms result in defect formation and phase transformation, which may cause significant changes of materials in composition and structure. These changes could modify or improve the various properties of materials. Therefore, The motivation for a large fraction of the work in this area has been to understand the effects of ion implantation on the microstructures and properties of materials. Many of the microstructure produced by implantation could be predicted from knowledge of radiation damage effects, by measuring the chemistry in the implanted layer, and referring to the equilibrium phase diagram[1]. Usually, ion implantation offers the possibility of obtaining alloys having: (1) a highly supersaturated solid solution with extended solid volubility over that predicted by the phase diagram; (2) a solid solution with stable and metastable phases; and (3) a disordered or amorphous structure[2]. Kujore et al[3] observed the defect structure of the ploycrystal copper implanted with Al, Cr and B ions by means of Transmission Electron Microscope (TEM). It was found that there existed dislocation loop and array on the surfaces of specimens implanted by various ions. The larger and the heavier of the ions are, the larger the defect density of the ion implanted microstructure will be. The microstructure of the surfaces implanted with chemical active ions (C, N, B, etc.) are usually composed of carbide, nitride, carbonitride and boride which come from the chemical reactions between substrate elements and ion implanted elements [4-6]. The implantation of metallic ions lead to the formation of the microstructure containing solid solution of the ion implanted atoms, intermetallic compounds formed between substrate atoms and metallic ions and carbides [7-9]. Ti+C dual ion implantation resulted in the formation of amorphous structure on many steel and alloy surfaces[1 O].

The objective of present work is to study the effect of W ion implantation on the microstructure defects and composition of pure iron by means of transmission electron microscopy (TEM) and Auger electron spectroscopy (AES) and X-ray photoelectron spectroscopy (XPS). The phase transformation and microstructural transition mechanisms were also briefly discussed.

2. Experimental details

The specimens used for implantation were pure iron in15mm×4.5mm×2.5mm flat pieces, polished with 1000 grit abrasive paper and cleaned ultrasonically in acetone. The implantation of tungsten ion into the specimen was conducted using a metal vapor vacuum arc (MEVVA) source ion implantation facility. The relevant implantation parameters are shown in Table 1.

Table 1. The ion implantation parameters of pure iron specimen

Ions	w+, W^{++}, w- , w- , W^{+++++}
Average charge state	2.74
Extraction voltage	48kV
Bulk temperature	≤150°C
Beam current	3mA
Dose (ions/cm*)	3×10^{17}

3 Experiment results and discussion

3.1. AES analysis

The ion implanted and other contamination elements on the surface of the ion implanted specimen were detected by AES on a Φ595 Scanning Auger Electron Spectroscope. Element distribution along depth was analyzed using AES with Ar^+ sputtering. The parameters of Ar^+ sputtering were 4keV in sputtering electrical voltage and 1 x $1mm^2$ in sputtering area. Fig. 1 shows the AES sputtering depth profile of the elements on the surface of W ion implanted pure iron, and it shows can be seen that tungsten ion were successfully implanted into pure iron specimen. Tungsten distributes in a plateau with an atomic concentration of 27at%, instead of Gaussian distribution, under the surface of the specimen. This is due to the fact that the ion implantation was conducted on a MEVVA source ion implanter under same acceleration voltage, and the implanted ions have several electrical charge states, thus, the ions with different charge states having deferent projected ranges. The plateau is the accumulating result of several Gaussian distribution with different projected ranges. On the ultra-shallow layer of the ion implanted specimen, oxygen has a highest atomic concentration, up to 42 at%, but, it decreases sharply with increase depth. It is the result of surface absorption and oxidation while specimen was exposed to air. Carbon has an approximate concentration of 16% from surface to certain depth in ion implanted layer. The surface carbon content may come from surface absorption of carbon-containing compounds after ion implantation. However, the carbon inside the implanted layer was incorporated by ion implantation through vacuum carbonization, ion recoil implantation and radiation enhanced diffusion. Iron only has an atomic concentration of 32at% on the surface. After that, it increases gradually with increase depth until it reaches steady stage.

3.2. XPS analysis

After ion implantation, the specimen surfaces were analyzed by XPS *on* a Φ550 Electron Spectrometer using Mg Kα line. The pass energy was 50 eV and the binding energy of C 1s (284.6eV) was used as a reference in XPS analysis. To get a thorough knowledge of the distribution of the chemical states, the implanted surfaces were sputtered by Ar ions under the conditions of acceleration voltage 3kV and sputtering area 5mm×5 mm. The durations of the sputtering time for different depths were O, 5, 15 and 30 minutes. Fig. 2 shows the XPS spectra of the W4f in the ion implanted pure iron specimen after different durations of sputtering. The results of analysis on the electron binding energies and possible chemical states of element in the first to fourth layers are summarized in Table 2.

From Fig. 2 and Table 2, it can be known that on the top surface of the ion implanted specimen $FeWO_4$ was the main existing state of the ion implanted element W. After 5 minute sputtering, the peak position of the binding energy shift to the low energy direction, which indicated that another oxide form of tungsten, WO_2, appeared. It was thought that all the oxides may be caused by the exposure of the ion implanted specimen to the ambient air after implantation. The contents of oxygen decrease with increase depth, as shown in Figure 1. It indicates that the amount

of oxides reduces as the depth increases. The atomic state of W becomes the dominant existing form of the implanted element after 15 minute sputtering. When the stage where oxygen are almost vanished and the incorporated carbon is at a high level of the content were reached after 30 minute sputtering, WC, carbide of W dominates the existing form of w.

Table 2 The electron binding energies and possible existing states of W in pure iron specimen implanted by W ion in dose of 3×10^{17} ions/cm^2.

Sputtering time (min.)	o	5	15	30
Binding energy (eV)	35.0	32.8	31.6	32.0
Existing state	$FeWO_4$	WO_2	w	WC

3.3 TEM observation of microstructure

Both the unimplanted and ion implanted specimens were cut off from the surface of the specimens in pieces of a thickness of 0.1 mm for TEM analysis. They were polished using abrasive paper until their thickness reach approximately 0.01 mm, and then were sputtered by Ar^+ ions to the TEM observable depth. The observation of the microstructure were performed on a H-600 Transmission Electron Microscope under the conditions of acceleration voltage of 100keV and instrument constant of $L\lambda$=29.6mmÅ.

The TEM observations of the unimplanted pure iron show that microstructure of it is composed of single phase et-Fe in size of 1 -4μm, with existing of lots of dislocations inside the grains. Figure 3 gives the TEM micrographs of pure iron implanted with W ion in dose of 3×10^{17} ions/cm^2, it can be known that microstructure of pure iron has been changed tremendously by ion implantation. It shows obvious irradiation damage caused by ion implantation, and no any trace of original grain boundary remains. The morphology of microstructure appears in an uneven island shape. Electron diffraction pattern indicates that it mainly consists of α-Fe. Further observation in a higher magnification (see Fig. 3b) indicates that each island is composed of many smaller islands. It gives out that W ion implantation caused the rupture of the α-Fe grains. In addition, The observation under more higher magnification (Figure 4) shows that there is some plain area between islands. The corresponding electron diffraction pattern indicates that these area has an amorphous structure. Therefore, it can be concluded that W ion implantation results in the rupture of the α-Fe grains and lattice distortion, and finally leads to the formation of amorphous phase in some area, with co-existing of the fine island-like α-Fe solid solution of W in the other area. Referring the AES and XPS analysis in 3.1 and 3.2, the amorphous phase should have a composition of Fe-W-C. Besides, there exists some big particles of Fe_3C in the ion implanted layer as observed in TEM analysis. The formation of Fe_3C phase is considered to be the result of carbon incorporation which has been discussed in AES analysis.

4. **Conclusion**

(1) The microstructure of W ion implanted pure iron consists of plain amorphous phase Fe-W-C, fine island-like et-Fe solid solution of W and particular carbide Fe_3C.

('2) The formation of the amorphous phase is the results of lattice distortion and grain rupture caused by ion collision and injection.

(3) The incorporation of carbon in implantation process resulted in the formation of Fe_3C.

5. **Reference:**

[1] D. I. Potter, M. Ashmed, and S. Lamond, Mat. Res. Sot. Symp. Proc., vol. 27(1 984), 117-126.
[2] K. V. Jata, D. ?Janoff, and E. A. Starke, Jr., Mat. Res. Sot. Symp. Proc., vol. 27(1984), 157-162.
[3] A. Kojore, S., B. Chakrabortty, E. A. Stroke and K. O. Legg, Nucl. Instru. Meth., 182/183 (1981 j 949-958.
[4] G. Longworht and N. E. W. Hartley, Thin Solid Films, 48(1978)95.
[5] V. P. Goltsev, V. V. Khodasevich, A. K. Kuleshov and V. V. Uglov, Nucl. Instr. Meth., B59/60 (1991) 751-756.
[6] D. Yang, X. Zhang, Q. Xue and H. Wang, Surface and Coatings Technology, 56 (1993), 119-124.
[7] X. G. Ning, C. Z. Ji and H. Q. Ye, Mat. Rec. Soc Symp. Proc., 235(1 992) 479.
[8] M. Kopcewicz, J. Jagielski, A. Turos and D. L. Williamson, Mat. Rec. Soc Symp. Proc., 235(1992) 509-514.
[9] T. Zhang, C. Z. Ji, J. Shen, J. Yang and J. Chen, Nucl. Instr. Meth., B59/60 (1991) 828-832.
[10] D. M. Follstaedt, F. G. Yost and L. E. Pope, Res. Sot. Symp. Proc., vol. 27 (1984), 655-660.

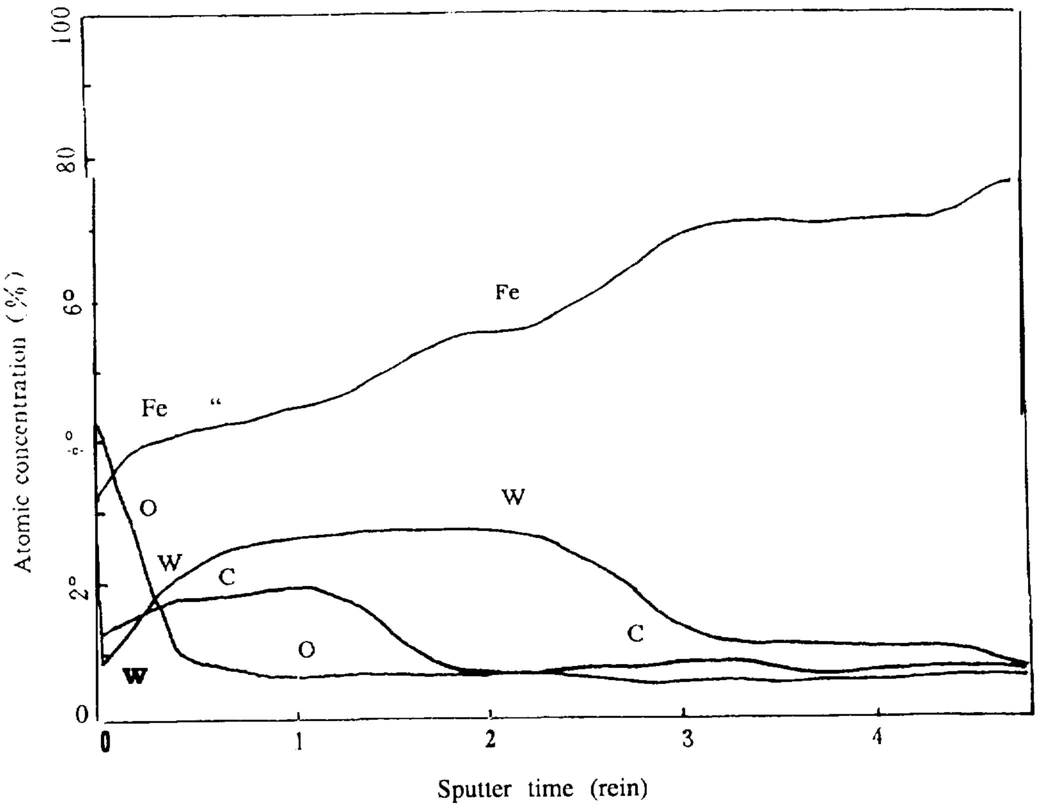


Fig. 1 AES sputtering depth profile of 3×10^{17} ions/cm^2 W ion implanted pure iron,

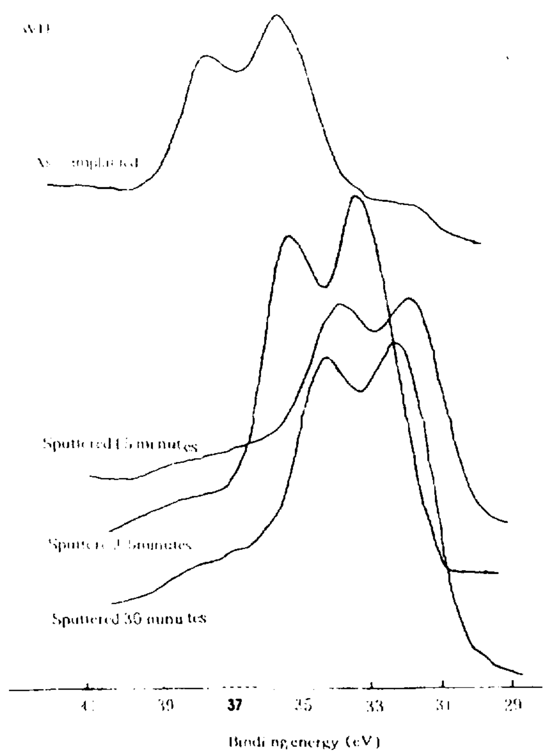
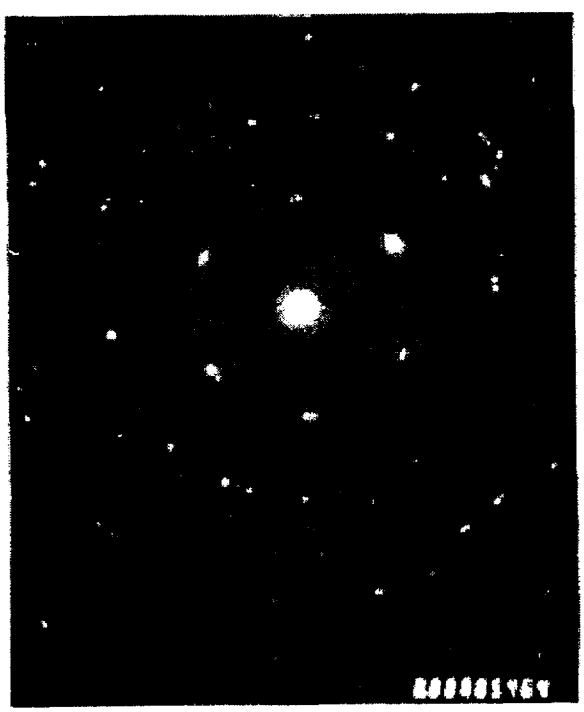


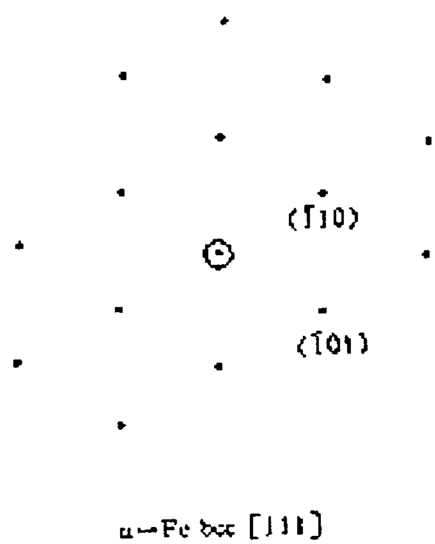
Figure 2. XPS spectra of W4f in pure iron specimen implanted by tungsten ions in dose of 3×10^{17} ions/cm^2 and sputtered by Ar ion for different durations of time.

a

b



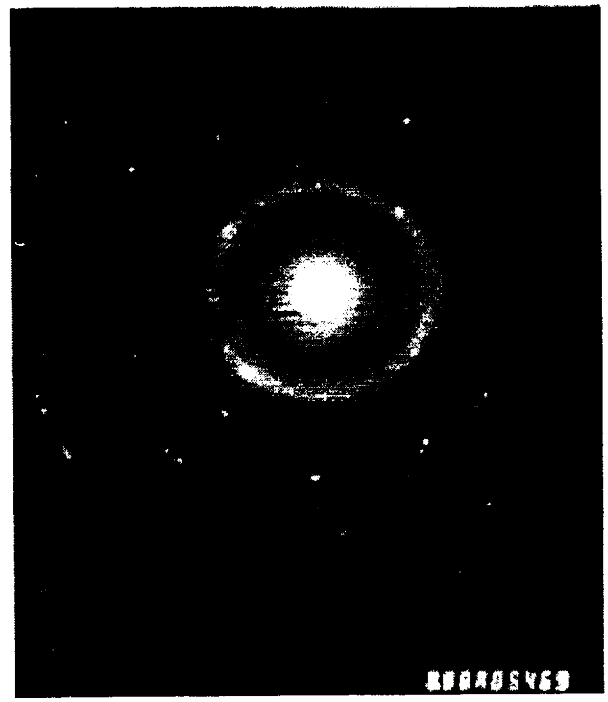
c



d

Fig. 3 TEM micrographs of 3×10^{17} ions/cm^2 W ion implanted pure iron showing ion implantation damage; a: bright-field, ×60k; b: bright-field, x 120k; c: selected area diffraction pattern; d: identification of Miller index and structure.

a



b

Fig. 4 TEM micrographs of 3×10^{17} ions/cm^2 W ion implanted pure iron showing coexistence of amorphous structure and crystal; a:bright-field, x 150k; b: selected area diffraction pattern.

URC97142

/ 141/63

Universal Approximation of Mamdani Fuzzy Controllers and Fuzzy Logical Controllers

Bo Yuan* **and George J.** Klir+

*NASA Center for Autonomous Control Engineering, Department of Engineering, New Mexico Highlands University. Las Vegas, New Mexico 87701

'Center for Intelligent Systems, Department of Systems Science and Industrial Engineering, Binghamton University, Binghamton, New York 13902

Abstract

In this paper, we first distinguish two types of fuzzy controllers, Mamdani fuzzy controllers and fuzzy logical controllers. Mamdani fuzzy controllers are based on the idea of int erpolat ion while fuzz y logical controllers are based on fuzzy logic in its narrow sense, i.e., fuzzy propositional logic. The two types of fuzzy controllers treat IF-THEN rules differently. In Mamdani fuzzy controllers, rules are treated disjunctively. In fuzzy logic controllers, rules are treated conjunctively. Finally, we provide a unified proof of the property of universal approximation for both types of fuzzy controllers.

1 Introduction

The study of universal approximation of a fuzzy controller was first init iated by Kosko [7], Wang [8] and Wang and Mendel [10]. It was an important contribution. to fuzzy control theory, since it provided a theoretical foundation for applying fuzzy controllers. It was shown that for a given continuous function defined on a compact domain, one always can design a fuzzy controller to approximate the function to any given precision, Kosko proved the result for his adaptive fuzzy system in [i'], Wang and Mendel provided a proof for a special case of Mamdani fuzzy cent rollers in [8] and [10]. Buckley proved the same result for Sugeno type fuzzy controllers in [2]. Ying presented a proof for a general Mamdani fuzzy controller in [13]. Castro provided another proof for a general Mamdani fuzzy controller in [3]. Klawonn and Novák proved the same result for fuzzy controllers based on fuzzy logical implications in [4].

Fuzzy controllers based on the idea of interpolation and those based on the idea of logical inference arc often not distinguished in the literature. More specifically, some t-norms, such as *min* and *product,* are often treated as fuzzy implications. The differences between the two types of fuzzy controllers are clearly stated in [4] and also can be found in [5]. In this paper, fuzzy controllers based on t-norms are called *Mamdani fuzzy* controllers, while *fuzzy-* controllers based on fuzzy implications arc called *fuzzy logical controllers.*

In Section 2, wc first review the structure of a fuzzy controller. The differences between Mamdani fuzzy controllers and fuzzy logical controllers are examined in Section 3. In Section 4, we present a unified proof of the property of universal approximation for both types of fuzzy cent rollers.

2 Fuzzy Controllers: A Brief Overview

'To build a fuzzy logic controller, one needs to follow the following four steps:

Step 1. Identifying state variables and control variables of a system to be controlled.

In this step, one has to determine relevant state and control variables, as well as the range of each oft hese variables. For inst ante, in the simple inverted pendulum example, the state variables are the angle of the pole, θ, and the rate of change of the angle, $\dot{\theta}$. The control variable is the force, f, applied to the cart carrying the pole. The range θ may, for example, be $(-\frac{\pi}{2}, \frac{\pi}{2})$.

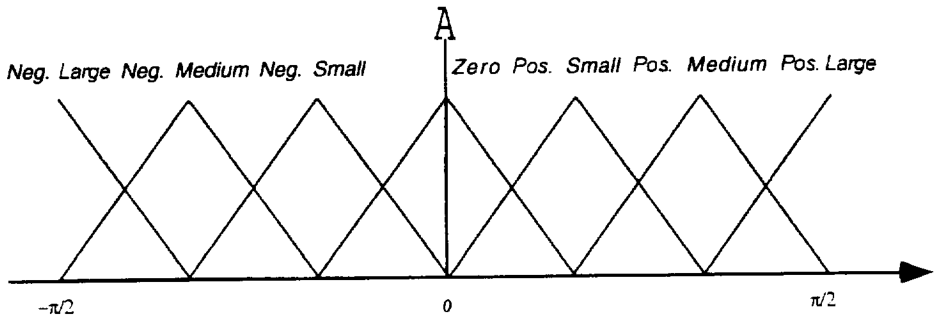


Figure 1: A fuzzy partition of the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$

Step 2. Dividing the range of each variable into several *level.*

The purpose of this step is to make it easier for domain experts to summarize their knowledge. The levels. which correspond to several linguistic terms, are used to describe the states of the system and cent ro: strategies. Usually, it is done by generating a fuzzy part it ion of the domain using fuzzy sets with triangular membership functions or, more generally, trapezoidal membership functions. Fuzzy sets with Gaussing membership functions, splines are also often used in many applications of neural-fuzzj' type control [I] and [9]. For example, let us *use* fuzzy sets with triangular membership functions to form. a fuzzy partition of the range of variable θ. We divide the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$ into seven levels, which arc labeled liguistically as negative large, negative *medium,* negative *small, approximately zero, positive small. positive median, positive large,* respectively (Fig. 1).

Step 3. Specifying IF-THEN inference rules for the controller.

These IF-THEN rules express domain experts' knowledge regarding the control task. Each rule describes a cent rol act ion that should be taken if the system is in one state. The general form of a rule in a rule base is

$$\text{If} \quad x \quad \text{is} \quad A \quad \text{then} \quad y \quad \text{is} \quad B$$

Intuitively, we should have IF-THEN rules that cover all possible states of the system that arc described by the linguistic terms generated in Step 2. Some researchers argue that fuzzy controllers wit h less IF-THEN rules have sufficiently good performance [12]. In these case, however, membership functions have to be carefully designed so that every state of the system is covered by some rules in the rule base.

Step 4. Selecting a defuzzification method.

A defuzzification method is a mapping $d : \mathcal{F}(\mathbb{R}) \to \mathbb{R}$, where $\mathcal{F}(\mathbb{R})$ denotes the fuzzy power set of $\mathbb{R}$. It maps a fuzzy set to a real number. There are many defuzzification methods in the literature. All defuzzification functions must satisfy the property

$$d(x) = X$$

for any $x \in \mathbb{R}$. One commonly used defuzzification method is called the *center of gravity method.* It takes the center of gravity of a fuzzy set as its defuzzification value. In this case, the fuzzy set is considered as the area surrounded by its membership function and the x-axis. This defuzzification method assumes that the membership function of the fuzzy set is integrable, which is true in most cases. Suppose $A(x)$ is the membership function of a fuzzy set A. Then the defuzzification value of fuzzy set A is

$$d(A) = \frac{\int x A(x) dx}{\int A(x) dx} \tag{1}$$

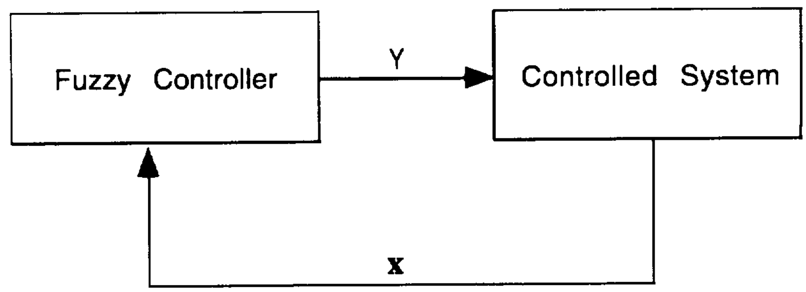


Figure 2: A general diagram of a fuzzy control system

for the continuous case, or

$$d(A) = \frac{\sum_{i=1}^{n} x_i A(x_i)}{\sum_{i=1}^{n} A(x_i)} \tag{2}$$

for the discrete case.

After one completes Step 1-4, he/she is ready to apply the fuzzy controller. A fuzzy controller generally has the structure show in Fig. 2. Given the current state of a controlled system or plant, fuzzy controller generates a value of the control variable. Then a control action is applied to the system, and the system, t bus, changes its state according to this control action. Therefore, a fuzzy cent roller can be considered as a function with the state of the system represented by its independent variables and the control state by its dependent variables. Suppose $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ denotes the n-dimensional state vector of the system, y denotes the control variable, then a fuzzy controller is a function

$$\begin{aligned} FS &: D \to \mathbb{R} \\ Y &= FS(\mathbf{x}) \end{aligned} \tag{3}$$

where $\Re$ is the real number set; D is a compact subset of n-dimensional space $\mathbb{R}^n$, which is specified in Step 1. Here. only one control variable is considered for the simplicity of discussion. For the case of several control variables. it can be easily decomposed into several fuzzy controllers with one cent rol variables.

The function FS is determined by the following procedure.

2.1 Fuzzification

At this stage, a fuzzy set is generated on the basis of current state vector $\mathbf{x}^t = (x_1^t, x_2^t, \ldots, x_n^t)^T$. There are many ways to generate this fuzzy set. One way is through generating symmetric triangular fuzzy numbers for all state variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$

$$A(x) = \bigwedge_{i=1}^{n} A_i(x_i) \tag{4}$$

where $\wedge$ denotes a t-norm [5] and A_i is determined by

$$A_i(x) = \begin{cases} 1 - \frac{|x - x_i^t|}{d_i} & \text{if } x \in [x_i^t - d_i, x_i^t + d_i] \\ 0 & \text{if otherwise;} \end{cases} \tag{5}$$

d_i is a constant which is monotonically related to the fuzziness of the fuzzification for each $i \in \{1,2,\ldots, n\}$. The larger d_i, the fuzzier the resulting fuzzy set A_i. Constants d_i **are** specified by the designer of the controller. In most recent applications, d_i is **assumed to** be zero for any $i \in \{1, 2, \cdots, n\}$. That is, this fuzzification step is ignored. [3] and [10].

2.2 Inference

At this stage, Zadeh's compositional rule of inference [6] and [11], is applied to calculate a fuzzy value of the control variable based on the fuzzy set obtained by fuzzification and the rule base specified in Step 3. The resulting fuzzy set B is calculated by

$$B(y) = \bigvee_{x \in D} i(A(x), R(x, y)), \tag{6}$$

where $\vee$ is a t-conorm and i is a t-norm [5]. R is a fuzzy relation defined on $D \times \Re$, which is determined by

$$R(x, y) = \diamond_{j=1}^{m} R_j(x, y), \tag{7}$$

where m is the number of rules in the rule base; o is either a t-norm or a t-conorm depending on the way R_j is calculated. R_j represents the jth rule in the rule base.

$$R_j(x, y) = f(A_j(x), B_j(y)), \tag{8}$$

where $f : [0, 1]^2 \to [0, 1]$ is a binary function that is either a t-norm or a *fuzzy* implication [.5]. When f is a t-norm. the fuzzy controller functions as interpolation; o must be a t-conorm. In this case, we call the fuzzy controller, FS. *a Mamdani fuzzy controller.* When f is a fuzzy implication, the fuzzy controller is based on logical inference, and o must be a t-norm. In this case, we call the fuzzy controller, FS. *a fuzzy logical controller* [4]. Here a fuzzy implication is required to satisfy at least the following two conditions

$$\begin{aligned} f(0,\ x) &= 1 \\ f(1,\ x) &= x \end{aligned} \tag{9}$$

for any $x \in [0, 1]$. Both R-implications and S-implications satisfy these conditions. For a detail discussion of fuzzy implications see [5].

2.3 Defuzzification

The fuzzy set B obtained by the inference described in Sec. 2.2 is defuzzified into a real number. The output of the fuzzy controller is the defuzzified value of That is

$$FS(x) = d(B).$$

3 Differences Between Mamdani Fuzzy Controllers and Fuzzy Logical Controllers

Let us consider a Mamdani fuzzy controller $FS_1 : D \to \mathbb{R}$, with $\vee = \diamond = \sup$, i= f= at-norm. According to Eq. (6)-(8), we have

$$\begin{aligned} B(y) &= \sup_{x \in D} i(A(x), R(x, y)) \\ &= \sup_{x \in D} i(A(x), \max_{j=1}^{m} i(A_j(x), B_j(y))) \\ &= \sup_{x \in D} \max_{j=1}^{m} i(A(x), i(A_j(x), B_j(y))) \\ &= \sup_{x \in D} \max_{j=1}^{m} i(i(A(x), A_j(x)), B_j(y)) \\ &= \max_{j=1}^{m} i(\sup_{x \in D} i(A(x), A_j(x)), B_j(y)) \end{aligned}$$

Therefore, for a Mamdani fuzzy controller, rules in the rule base are considered disjunctively. One can apply each rule in a rule base to generate a result, then take the maximum of all results as the resulting fuzzy set.

Now, let us consider a **fuzzy** logical controller $FS_2 : D \rightarrow \mathbb{R}$. with V = SUP, o = min, i = a t-norm and f = a fuzzy implication. According to Eq. (6)-(8), we have

$$B(y) = \sup_{x \in D} i(A(x), R(x, y)) = \sup_{x \in D} i(A(x), \min_{j=1}^{m} f(A_j(x), B_j(y))) \tag{10}$$

Suppose A is a crisp point x_0 in $\mathbb{R}^n$, that is,

$$A(x) = \begin{cases} 1 & \text{if } x = x_0 \\ 0 & \text{if } x \neq x_0 \end{cases}$$

Then, Eq. (10) becomes

$$B(y) = \min_{j=1}^{m} f(A_j(x_0), B_j(y))) \tag{11}$$

Therefore, we can see that a fuzzy logical controller treats rules in a rule base conjunctively. In the case of no fuzzification, one applies each rule in the rule base and takes then the minimum of all obtained results as the final resulting fuzzy set.

In conclusion, we can see that one major difference between a Mamdani fuzzy controller and a fuzzy logical controller is that the former treats rules in a rule base disjunctively while the latter treats them conjunctively. For a more detail discussion of the differences between Mamdani fuzzy controllers and fuzzy logical controllers, see Ref. [4].

4 Universal Approximation of Mamdani and Fuzzy Logical Controllers

Wang and Mendel in [10] have shown that Mamdani fuzzy controllers with f = *product* are universal approximators. Ying [13] and Castro [3] have proven this for a general Mamdani fuzzy controller, using different approaches. Klawonn and Novák in [4] proved that fuzzy logical controllers are universal approximators. Here, we show a unified proof of the same result for both types of fuzzy controllers.

Theorem 1 *Let $g : U \rightarrow \mathbb{R}$ be a continuous function from a compact subset U of $\mathbb{R}^n$ to $\mathbb{R}$. Then for any $\varepsilon > 0$, we always can find a fuzzy controller FS such that*

$$|FS(x) - g(x)| < \varepsilon \tag{12}$$

for any $x \in U$.

Proof. Since g is continuous, for any $\varepsilon > 0$, and any $u \in U$, there exists $d(u, \varepsilon) > 0$, such that for any $x \in B(u, d(u, \varepsilon)) = \{x \in U \,|\, |x - u| < d(u, \varepsilon)\}$,

$$|g(x) - g(u)| < \varepsilon. \tag{13}$$

Since $U \subseteq \bigcup_{u \in U} B(u, d(u, \varepsilon))$ and U is a compact set, there exists a finite number of elements $\{u_1, u_2, \cdots, u_m\}$ such that

$$U \subseteq \bigcup_{i=1}^{m} B(u_i, d(u_i, \varepsilon)) \tag{14}$$

Now we generate a finite disjoint cover of U based on Eq. (14). Let $A_1 = B(u_1, d(u_1, \varepsilon))$,

$$A_i = B(u_i, d(u_i, \varepsilon)) \setminus \bigcup_{j=1}^{i-1} A_j \tag{15}$$

for $i = 2, 3, \cdots, m$. It is easy to see that

$$A_i \cap A_j = \emptyset$$

for $i \neq j$, and

$$U \subseteq \bigcup_{i=1}^{m} A_i \tag{16}$$

Now, we build a fuzzy controller without the fuzzification step and based on the following inference rules in the rule base

$$\text{If x is } Al \text{ then } y \text{ is } B_1 = g(\mathbf{u}_1)$$
$$\cdots\cdots$$
$$\text{If x is } A_m \textbf{ then } y \textbf{ is } B_m = g(\mathbf{u}_m)$$

For any $x \in U$, it follows from Eq. (16) that there exists a unique $i_0 \in \{1, 2, \ldots, m\}$ such that $x \in A_{i_0}$ and $x \notin A_i$ for $i \neq i_0$. That is

$$A_{i_0}(\mathrm{x}) = 1 \text{ and } A_i(\mathrm{x}) = O \text{ for } i \neq i_0.$$

According to Eq. (6)-(8), the resulting fuzzy set is

$$\begin{aligned} B(y) &= \bigvee_{\mathrm{x}' \in D} i(A(\mathrm{x}'), R(\mathrm{x}', y)) \\ &= R(\mathrm{x}, \mathrm{y})) \\ &= \circ_{i=1}^{m} R_i(\mathrm{x}, \mathrm{y})" \\ &= \circ_{i=1}^{m} f(A_i(\mathrm{x}), B_i(y)) \end{aligned}$$

for any $y \in \mathbb{R}$.

When f is a t-norm and o is a t-conorm,

$$\begin{aligned} B(y) &= f(1, B_{i_0}(y)) \\ &= B_{i_0}(y) \end{aligned}$$

Therefore

$$FS(\mathrm{x}) = d(B) = d(B_{i_0}) = g(\mathbf{u}_{i_0}).$$

for any type of defuzzification function d.

When f is a fuzzy implication and o is a t-norm

$$\begin{aligned} B(y) &= \circ(\circ_{i \neq i_0} f(A_i(\mathrm{x}), B_i(y)), f(A_{i_0}(x)> B_{i_0}(Y))) \\ &= \circ(\circ_{i \neq i_0} f(O, B_i(y)), f(A_{i_0}(\mathrm{x}), B_{i_0}(y))) \\ &= \circ(1, f(A_{i_0}(\mathrm{x}), B_{i_0}(y))) \\ &= f(A_{i_0}(x)> B_{i_0}(y)) \\ &= f(1, B_{i_0}(y)) \\ &= B_{i_0}(y) \end{aligned}$$

Therefore, again we have

$$FS(\mathrm{x}) = \mathrm{d(B)} = d(B_{i_0}) = g(\mathbf{u}_{i_0}).$$

for any type of defuzzification function d.

Finally, according to Eq. (13), we have

$$|FS(\mathrm{x}) \ g(\mathrm{x})| \ \ |g(\mathbf{u}_{i_0}) \ g(\mathrm{x})| < \varepsilon.$$

This completes the proof.

5 Conclusions

In this paper, we argue that two different types of fuzzy controllers, Mamdani fuzzy controllers and fuzzy logical controllers, should be distinguished. We also present a unified proof of the property of universal approximation for these two types of fuzzy controllers.

6 Acknowledgment

This research has been sponsored by NASA Center for Autonomous Control Engineering. The first author would like to thank Dr. Djuro G. Zrilic for his support of this research..

References

[1] M. Brown and C. Harris. *Neurofuzzy adaptive modelling and control.* Prentice Hall, New York, 1994.

[2] J. J. Buckley. Sugeno type controllers are universal controllers. *Fuzzy Sets* and Systems, 53(3):299–304, 1993.

[3] J. Castro. Fuzzy logic controllers are universal approximators. *IEEE Trans. on Systems, Man and Cybernetics, 25(4)* :629–635, *1995.*

[4] F. Klawonn and V. Novak. The relation between inference and interpolation in the framework of fuzzy systems. *Fuzzy Sets and Systems, 81(3)* :331–354, *1996.*

[5] G. J. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic: Theory and Applications.* Prentice Hall PTR, Upper Saddle River, NJ, 1995.

[6] G. J. Klir and B. Yuan, editors. *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh.* World Scientific, Singapore, 1996.

[7] B. Kosko. Fuzzy systems as universal approximators. *Proc. of First IEEE Intern. Conference on Fuzzy Systems,* II: 1153–1162, 1992.

[8] L. X. Wang. Fuzzy systems are universal approximators. *Proc. of First IEEE Intern. Conference on Fuzzy Systems,* II: *1163-1170, 1992.*

[9] L. X. Wang. *Adaptive Fuzzy Systems and Control: Design and Stability Analysis.* Prentice Hall, Englewood Cliffs, NJ, 1994.

[10] L. X. Wang and J. M. Mendel. Fuzzy basis functions, universal approximation, and orthogonal least-squares learning. *IEEE Trans. on Neural Networks, 3(5):807–814, 1992.*

[11] R. R. Yager, S. Ovchinnikov, R. M. Tong, and H. T. Nguyen, editors. *Fuzzy Sets and Applications: Selected Papers by L. A. Zadeh.* Wiley - Interscience, New York, 1987.

[12] T. Yamakawa. A fuzzy inference engine in nonlinear analog mode and its application to a fuzzy loge control. *IEEE Trans. on Neural Networks,* 4(3):496–522, *1993.*

[13] H. Ying. Sufficient conditions on general fuzzy systems as function approximators. *Automatica.* 30(3):521–525, *1994.*

Page intentionally left blank

URC97143

0142/76

Ion Beam Synthesis and Optical Properties of Semiconductor Nanocrystals and Quantum Dots

Jane G. Zhu, * C. W. White, S. P. Withrow, and J. D. Budai
Oak Ridge National Laboratory, Solid State Division
Oak Ridge, TN 37831-6057

D. O. Henderson
Fisk University, Physics Department
Nashville, TN 37208

ABSTRACT

Nanocrystals of semiconductor materials have been fabricated in SiO_2 by ion implantation and subsequent thermal annealing. Strong red photoluminescence (PL) peaked around 750 run has been observed in samples containing Si nanocrystals in SiO_2. The Si nanocrystals in the samples with optimized PL intensities are a few nanometers in diameter. Difference in the absorption bandgap energies and the PL peak energies are discussed. Significant influence of implantation sequence on the formation of compound semiconductor nanocrystals are demonstrated with the GaAs in the SiO_2 system. Optical absorption measurements show that Ga particles have already formed in the as-implanted stage if Ga is implanted frost. A single surface phonon mode has been observed in the infrared reflectance measurement from samples containing GaAs nanocrystals.

INTRODUCTION

Ion implantation is a very useful technique for altering the near-surface properties of a wide range of materials. Desired elements can be injected into a solid in a controlled and reproducible manner by ion implantation [1]. Supersaturated impurity concentrations can be produced by high-dose implantation. Subsequent annealing leads to precipitation and the formation of nanocrystals which are encapsulated in the host material. Semiconductor nanocrystals and quantum dots have attracted considerable interest due to potential applications in novel optoelectronic devices [2]. Quantum confinement effects are expected when the sizes of the nanocrystals are smaller than the exciton diameters. For optical applications, for example, visible luminescence can be obtained from group IV nanocrystals due to quantum confinement effects [3]. Nonlinear optical properties can also be significantly enhanced for the nanocrystals compared to the bulk material [4].

We have studied the formation and properties of elemental (group IV) and compound semiconductor (III-V and II-VI) nanocrystals in different host materials [5-8]. The nanocrystal sizes can be controlled by the ion implantation dose and annealing temperatures. To form compound semiconductor nanocrystals, the constituent elements are implanted sequentially. Examples reported in this paper include Si and GaAs nanocrystals formed in amorphous SiO_2 matrices.

EXPERIMENTAL PROCEDURES

The semiconductor nanocrystals were formed by ion implantation of appropriate semiconductor species into a SiO_2 layer on (100) silicon substrate, followed by subsequent thermal annealing. The SiO_2 layer was -750 nm thick, formed by thermally oxidizing a (100) Si wafer. Samples were also prepared by ion implantation into fused silica substrates (Coming 7940). Ion implantation was done at room temperature with doses in the range of $(3\text{-}30)\times10^{16}$

•Permanent address: Department of Physics, New Mexico State University, Box 30001, Dept. 3D, Las Cruces, NM 88003-8001.

ions/cm^2. The implant energies were chosen to put the peak concentration at the middle of the oxide layer. Samples were annealed isochronall y for 1 h under Ar + 4%H_2 ambient at atmospheric pressure. The annealing temperatures were within the range of 600°C to 1100°C. The nanocrystalline structures were investigated by transmission electron microscopy (TEM) and X-ray diffraction. All the TEM specimens were prepared in cross sections, since the concentration distribution from ion implantation is a function of depth. Depth profiles were also examined by Rutherford backscattering spectrometry (RBS). Optical measurements include optical absorption, photoluminescence (PL) and infrared reflectance measurements.

Si NANOCRYSTALS IN SiO_2

Multiple implants of Si at different energies were performed to produce uniform concentration profiles of Si inside SiO_2. A TEM micrograph in Fig. 1 shows the Si nanocrystals are a few nanometers in size (in the 1-5 nm range, mostly -2-3 nm) in the samples implanted with excess Si concentration of -5 x 10^{21} cm^{-3} throughout a SiO_2 film on a Si substrate and annealed at 1100°C for 1 h. The formation of nanocrystals of different materials can be very different under similar annealing conditions [6]. It takes a higher annealing temperature for Si than that, e.g., for Ge to precipitate inside SiO_2. To achieve visible luminescence, the sizes of Si nanocrystals should be about a few nanometers in order to have sufficient quantum confinement effect [3].

Very strong PL has been observed in the samples containing Si nanocrystals about a few nanometers in size. Figure 2(a) shows PL spectra, excited by an Ar *laser at* 514.5 nm wavelength, from a samples implanted with Si at a dose of 1.5 x 10^{17} cm^{-2} before and after thermal annealing. There is some luminescence with a broad peak centered at -650 nm from the as-implanted sample. These luminescence centers are presumably due to the ion-implantation damage of the SiO_2 matrix. The PL intensity is significantly reduced after the sample is annealed at 800°C. After annealing at a higher temperature of 1100°C, Si nanocrystals in sizes of a few nanometers have been formed and a very strong PL peak center at -750 nm have been observed. For samples annealed at different temperatures, the PL peak intensity is the highest for the sample annealed at 1100°C.

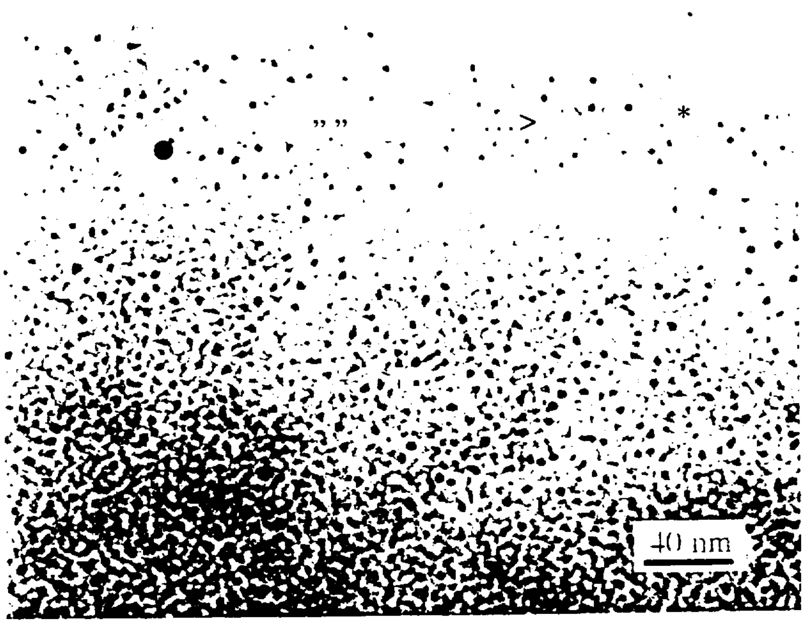


Fig. 1. TEM micrograph showing the Si nanocrystals formed inside a SiO_2 matrix.

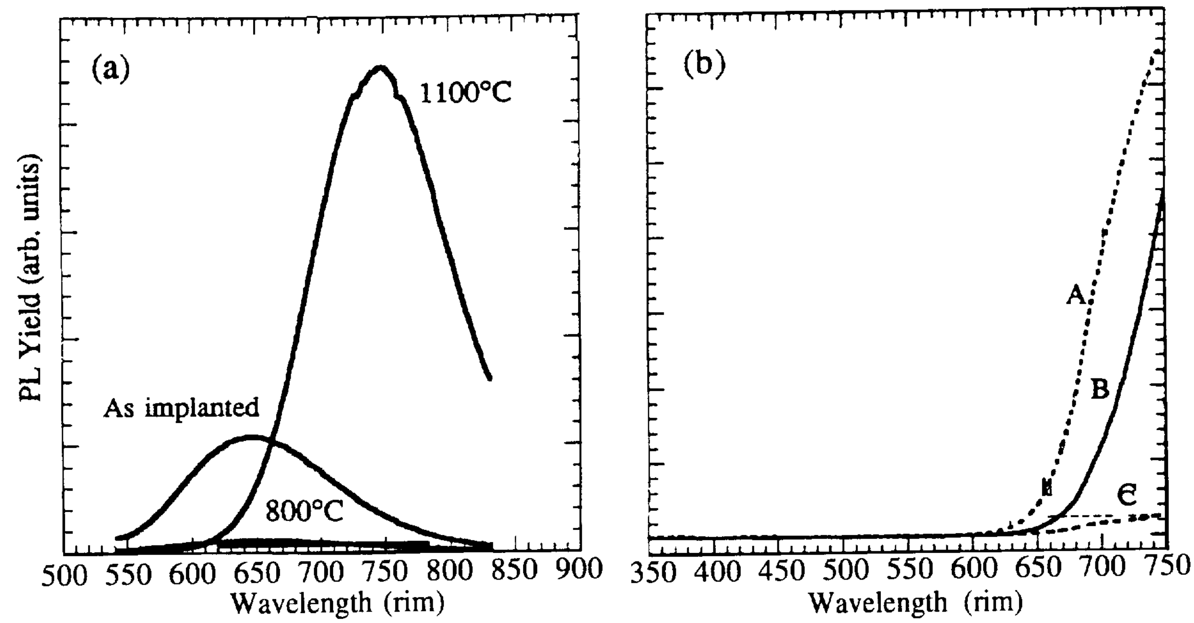


Fig. 2. (a) PL spectra from a sample implanted with Si at a dose of 1.5×10^{17} cm^{-2} before and after thermal annealing at 800°C and 1100°C. PL was excited by an Ar laser at 514.5 nm wavelength. (b) PL spectra from samples implanted with flat Si concentration profiles of A: 1×10^{21} cm^{-3}, B: 5×10^{21} cm^{-3}, and C: 2×10^{22} cm^{-3}. These spectra were excited by a Xe lamp at 300 nm wavelength.

Figure 2(b) shows PL spectra, measured using a SPEX Fluorolog 2 and excited at 300 nm, from samples implanted with flat Si concentration profiles of (A) 1×10^{21} cm^{-3}, (B) 5×10^{21} cm^{-3}, and (C) 2×10^{22} cm^{-3} and subsequently annealed at 1100°C (the detector's detection limit does not cover the whole range of the PL peaks). The only strong PL peak occurs around 750 nm. When measured using an Ar laser, excited at 514.5 nm, the PL peak from sample B has the highest intensity (Si nanoparticles in this sample are shown in Fig. 1). The PL peak is shifted toward a higher energy for sample A, which has been implanted with less Si, but shifted to a lower energy with much lower peak intensity for sample C which has much more Si implanted TEM of sample C revealed that the Si nanocrystals are in the range from a few to several nanometers. Combining the results from PL and TEM from a number of samples, it is concluded that maximum PL intensity occurs in the samples annealed at 1100°C and containing Si nanocrystals with sizes about a few nanometers.

Strong optical absorption has been observed in the Si-nanocrystal-containing silica samples. The absorption bandgap energies measured are much larger than the PL peak energies observed [9]. When the implantation dose of Si is changed, the size distributions of Si nanocrystals change and the strong peak position can be shifted slightly, much less than the shift in the absorption energy. Large differences between luminescence and optical absorption energies have been reported for porous Si produced by electrochemical dissolution of Si [10]. Similarly, the optical properties of these Si-nanocrystals-in-SiO_2 systems can be explained by quantum confinement effects in Si nanocrystallites in association with surface/interface states. Theoretical calculations have also been reported to show the existence of self-trapped excitons on bonds of silicon crystallite [11]. The self-trapped state can be a stable situation for the excited states for very small crystallite.

For the formation of compound semiconductors, such as GaAs, where there is more than one element involved, the sequence of Ga and As implantation was found to influence the size distributions of GaAs nanocrystals dramatically [12]. Figure 3 shows cross-sectional TEM images from samples implanted with the same doses, 1.5 x 10^{17} ions/cm^2, but different implantation sequences and annealed at 1000°C for 1 h. The GaAs nanocrystals in Fig. 3(a) were formed in the sample implanted with Ga first and then As, and have sizes ranging from a few nanometers to -30 nm. These nanocrystals are nearly spherical and randomly oriented with respect to each other, as expected for amorphous matrices. Some voids are observed in the region near the oxide surface. In the sample implanted with As and then Ga, the GaAs nanocrystals formed after annealing at the same temperature are much smaller with sizes in the range of 1- 10 nm as shown in Fig. 3(b). The difference between the two samples shown in Fig. 3 (a) and (b) is very striking_considering they have been through the same processing except the implantation sequence of Ga and As.

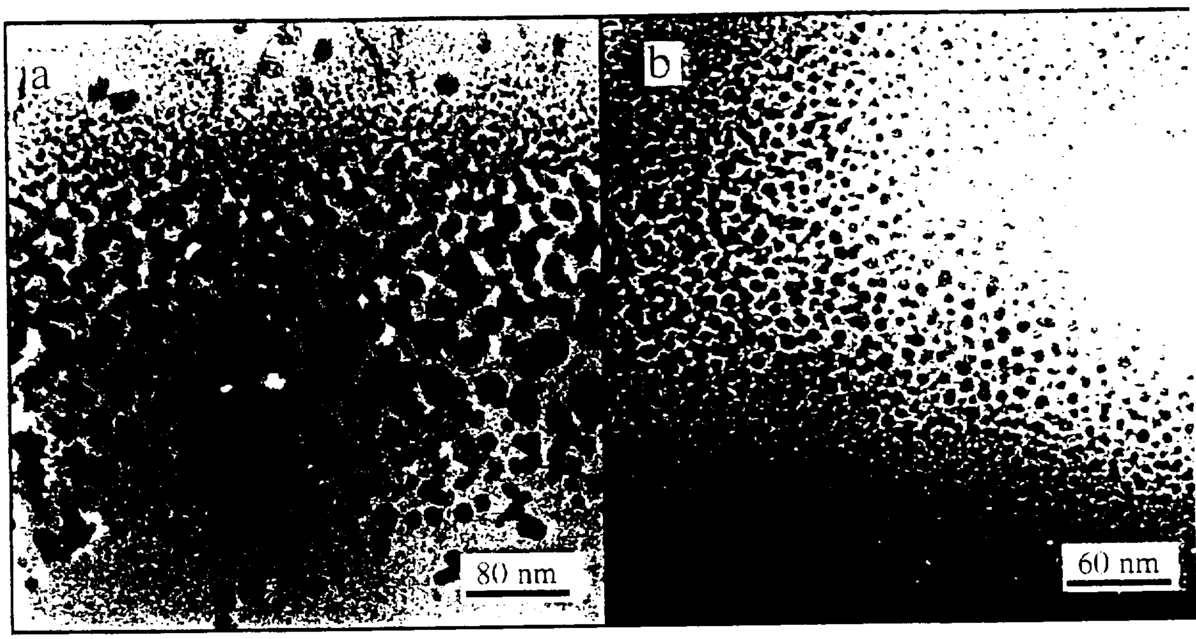


Fig. 3. Cross-sectional TEM images from samples implanted with equal doses, 1.5 x 10^{17}cm^{-2}, of Ga and As, but different implantation sequence: (a) Ga first and then As and (b) As first and then Ga. Both samples have been annealed at 1000°C.

RBS measurements of the implanted ion profiles reveal that the As implanted in SiO_2 is thermally stable, while Ga implanted in SiO_2 is very mobile during implantation and annealing. Optical absorption spectra from samples implanted with Ga along or As along into silica glass wafers are shown in Fig. 4. A strong absorption peak at 220 nm has been observed in the as-implanted sample implanted with Ga only. This absorption peak is attributed to the surface plasmon resonance of Ga particles in SiO_2. Details on surface plasmon resonance of metal particles, such as Ag, in a dielectric matrix can be found elsewhere [13]. The optical density decreases after annealing at 1000°C, which is in agreement with the Ga loss observed in the RBS spectra. TEM from a sample implanted with Ga first and then As confirms the formation of Ga particles in the as-implanted stage. The optical absorption spectrum for the sample implanted with As only is virtually unchanged after annealing at 1000°C. The thermal stability of the As profile helps to interpret the thermal stability of the ion-concentration profile of As + Ga in the samples implanted with As fret. When Ga ions are implanted after the implantation of As, they bond with

the As atoms due to the chemical affinity between Ga and As. In the samples implanted with Ga first and then As, there is much more diffusion involved since Ga is very mobile in SiO_2. Consequently, the GaAs nanocrystals grow much bigger.

In the samples implanted with As first and then Ga, the GaAs nanocrystals formed after annealing are smaller than the bulk exciton radius. Therefore, quantum confinement effect is expected. An infrared reflectance spectrum is shown in Fig. 5 recorded from a sample implanted with equal amount, 1.0×10^{17} ions/cm^2, of As and then Ga and annealed at 1000°C for 1 h. A strong reflectance at 278 cm^{-1} has been observed. A Raman spectrum measured from the bulk GaAs is also plotted in Fig. 5. The single peak from the GaAs-nanocrystal-containing sample lies in between the transverse (267 cm^{-1}) and longitudinal (291 cm^{-1}) optical phonon modes of bulk GaAs. Observation of a similar single mode between the bulk transverse and longitudinal optical modes in the infrared spectrum has been reported for microcrystallites of MgO and GaP [14]. This peak is attributed to the excitation of surface phonon modes. A single surface phonon mode has also been detected for GaAs nanocrystals formed in single crystalline Al_2O_3 matrix, where GaAs nanocrystals are faceted and very well aligned with the crystalline matrix [7].

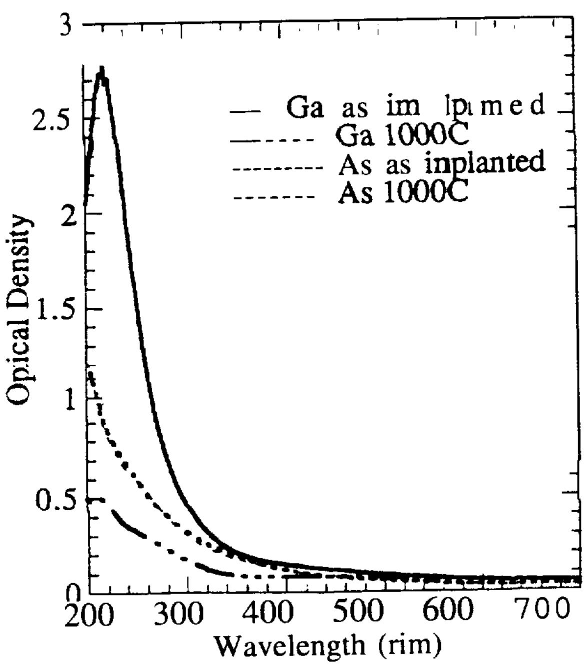


Fig. 4. Optical absorption spectra from samples implanted with 1.0×10^{17}/cm^2 of Ga or As only, before and after annealing.

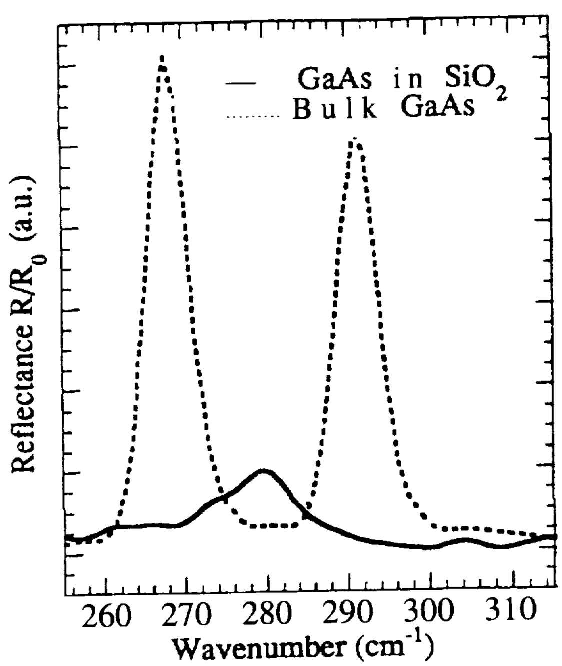


Fig. 5. Infrared reflectance spectrum (solid curve) from a GaAs-nanocrystal-containing sample with the Raman spectrum (dashed curve) from bulk GaAs superimposed.

CONCLUSIONS

Semiconductor nanocrystals of Si and GaAs have been formed in SiO_2 by ion implantation and subsequent thermal annealing. The microstructure has been characterized extensively by TEM. A broad range of nanocrystal sizes can be produced through the control of ion implantation and annealing processes. Strong PL peaked at around 750 nm is observed in Si-nanocrystal-containing samples. The absorption bandgap energies measured are consistent with the quantum confinement effect, while the PL peak energies are considered to be associated with the surface/interface states. For the formation of compound semiconductor GaAs nanocrystals, it is demonstrated that the sequence of Ga and As ion implantation affects the size distributions of GaAs nanocrystals significantly. The nanocrystal sizes are much bigger in the samples with Ga

implanted first than those with As implanted first. This phenomenon is explained by the different diffusion behaviors of Ga and As species. Optical absorption measurements show that Ga particles have already formed in the as-implanted stage with the absorption peak at 220 nm due to a surface plasmon resonance of metal particles in SiO_2. Single surface phonon mode has been observed from samples containing GaAs nanocrystals in. the infrared reflectance measurement. We have also fabricated and characterized nanocrystals of other semiconductor materials in different matrices using the ion implantation technique, which is beyond the scope of this paper.

ACKNOWLEDGMENTS

This research was sponsored by the Division of Materials Sciences, U.S. Department of Energy, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp. and contract DE-FG05-94ER45521 with Fisk University, and supported in part through the Oak Ridge National Laboratory Postdoctoral Research Program administered by Oak Ridge Institute for Science and Technology.

REFERENCES

1. C. W. White, C. J. McHargue, P. S. Sklad, L. A. Boatner, and G. C. Farlow, Mater. Sci. Reports 4,41 (1989).
2. See papers included in *Microcrystalline and Nanocrystalline Semiconductors,* edited by R. W. Collins, C. C. Tsai, M. Hirose, F. Koch, and L. Brus, Mater. Res. SW. Symp. Proc. 358, (MRS, Pittsburgh, 1995).
3. T. Takagahara and K. Takeda, Phys. Rev. B 46, 15578 (1992).
4. L. Brus, Appl. Phys. A 53, 465 (1991).
5. J. G. Zhu, C. W. White, J. D. Budai, S. P. Withrow, and Y. Chen, Mater. Res. Sot. Symp. Proc. 358, 175 (1995).
6. J. G. Zhu, C. W. White, J. D. Budai, S. P. Withrow, and Y. Chen, J. Appl. Phys. 78, 4386 (1995).
7. C. W. White, J. D. Budai, J. G. Zhu, S. P. Withrow, D. H. Hembree, D. O. Henderson, A. Ueda, Y. S. Tung, R. Mu, and R. H. Magruder, J. Appl. Phys. 79, 1876 (1996).
8. C. W. White, J. D. Budai, J. G. Zhu, S. P. Withrow, and M. J. Aziz, Appl. Phys. Lett. 68, 2389 (1996).
9. C. W. White, J. D. Budai, S. P. Withrow, J. G. Zhu, S. J. Pennycook, R. A. Zuhr, D. M. Hembree Jr., and D. O. Henderson, R. H. Magruder, M. J. Yacaman, G. Mondragon, and S. Prawer, paper to be published in Nucl. Instr. Methods B.
10. D. J. Lockwood, Solid State Commun. 92, 101 (1994).
11. G. Allan, C. Delerue, and M. Lannoo, to be published.
12. J. G. Zhu, C. W. White, D. J. Wallis, J. D. Budai, and S. P. Withrow, and D. O. Henderson, Mat. Res. Sot. Symp. Proc. 396, 447 (1996).
13. G. W. Arnold and J. A. Borders, J. Appl. Phys. 48, 1488 (1977).
14. S. Hayashi, Jpn. J. Appl. Phys. 23, 665 (1984).

5143/63

URC97144

Various Hardware Implementations of Membership Function Generators

Djuro G. Zrilic+, Jaime Ramírez-Angulo*, and Bo Yuan+

'NASA Center for Autonomous Control Engineering, Department of Engineering, New Mexico Highlands University, Las Vegas, New Mexico 87701

*Klipsch School of Electrical and Computer Engineering, New Mexico State University, Las Cruces, New Mexico 88003

Abstract

The choice of fuzzy set membership functions affects how well fuzzy systems approximate functions. The most common types of fuzzy membership functions are: triangular, trapezoids and Gaussion functions. In this paper, we propose a unified solution in hardware implementations of a wide variety of membership functions of fuzzy set. This provides designers more flexibility in designing fuzzy systems.

1 Introduction

There exist numerous examples of successful applications of fuzzy logic in control, pattern recog-nit ion, and expert systems. Fuzzy logic has been applied into more and more scientific areas. The design of fuzzy systems is very much depended on knowledge of domain experts. Due to high degree of freedom in designing a fuzzy logic system, fine tuning or optimizing of a fuzzy system is necessary.

Majority of fuzzy logic applications are built using existing commercial fuzzy logic processors. These applications usually do not require high speed performance. There are some applications, however, in which processing speed is very important.

Our goal, which is presented in this paper, is to design " parallel" and "serial" membership function generators (MFG) with programmable parameters such as, *horizontal shift, shape selection, height specification,* etc. During last 12 months, we proposed several different methods to implement membership function generators in hardware. We use some basic fuzzy sets defined on the unit interval [0, 1] to construct different shapes of membership functions. Those basic fuzzy sets on [0, 1] corresponds to some basic linguistic truth values such as *true, very true, fairly true, false, very false, fairly false.*

2 Theoretical Background

The general block diagram of our membership function generator is shown in Fig. 1. The system consists three blocks: pre-processor, shape generator and post-processor or scaler.

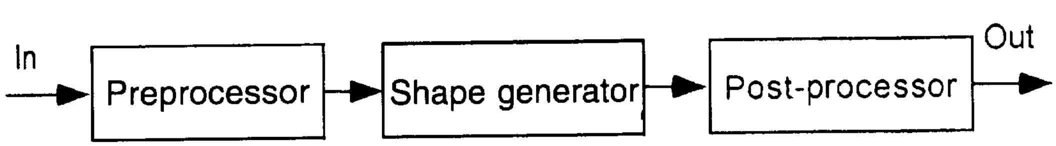


Figure 1: The general block diagram of membership function generator circuits

The role of the preprocessor is to determine a proper region in which a current input voltage locates. It is done by comparing the input voltage with programmed parameters. The comparison can be done in either digital or analog manner. The role of the shape generator is to generate a selected shape of the membership function. Shapes can be linear, which is a straight line, or nonlinear, which can be any desired shape. The hardware of the shape generator depends on the

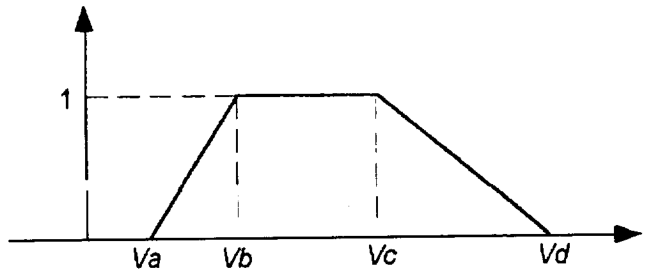


Figure 2: A trapezoidal membership function

specific type of membership functions. When only linear membership functions are involved, only a simple analog circuit is sufficient. When nonlinear membership functions are involved, we use either the RAM approach for high speed parallel processing or dedicated analog and digital circuit to implement some typical types of nonlinear membership functions.

In applications of fuzzy sets and fuzzy logic, trapezoidal membership functions are the most often used. The major reason is that a trapezoidal membership function is a piecewise linear function, which is easy to be handled theoretically and easy to be implemented in hardware. A trapezoidal membership function *A(v)* can have the following general form:

$$A(v) = \begin{cases} 0 & \text{if } v < V_a \\ \frac{v-V_a}{V_b-V_a} & \text{if } V_a \leq v < V_b \\ 1 & \text{if } V_b \leq v \leq V_c \\ \frac{V_d-v}{V_d-V_c} & \text{if } V_c < v \leq V_d \\ 0 & \text{if } V_d < v \end{cases} \tag{1}$$

where $V_a, V_b, V_=, V_d$ are parameters that fully specify the shape of the trapezoidal shown in Fig. 2.

Trapezoidal membership functions are special cases of so-called fuzzy numbers. Fuzzy numbers arc fuzzy sets defined on the real number set that are convex and with height 1. See [2]. A fuzzy number A has the following general form:

$$A(v) = \begin{cases} 0 & \text{if } v < V_a \\ l(v) & \text{if } V_a \leq v < V_b \\ 1 & \text{if } V_b \leq v \leq V_c \\ r(v) & \text{if } V_c < v \leq V_d \\ o & \text{if } V_d < v \end{cases} \tag{2}$$

where $l(v)$ is called the *left function* of the fuzzy number which is increasing and cent inuous from right; and $r(v)$ is called the right function which is decreasing and continuous from left. Interval $[V_b, V_c]$ is called the core of the fuzzy number. From Eq. 1, we can see that the trapezoidal membership function has $l(v) = \frac{v-V_a}{V_b-V_a}$ and $r(v) = \frac{V_d-v}{V_d-V_c}$.

To implement a general fuzzy number in hardware, the difficulty is the left and right functions in Eq. 2. It is impossible to have a hardware that can realize all possible functions $l(v)$ and r(v). In this paper, we propose to implement some nonlinear functions via using some typical functions defined on the unit interval. The two typical functions on the unit interval are particular interesting to us are the square function and the square root function that correspond to linguistic truth values *very true* and *fairly true*, respectively. The two functions, v^2, and $\sqrt{v}$, are depicted in Fig. 3.

In our implementation, the left function, $l(v)$, and the right function, r(v), of a fuzzy number are the composition of functions in the unit interval and the straight line. That is,

$$l(v) = f\left(\frac{v - V_a}{V_b - V_a}\right)$$

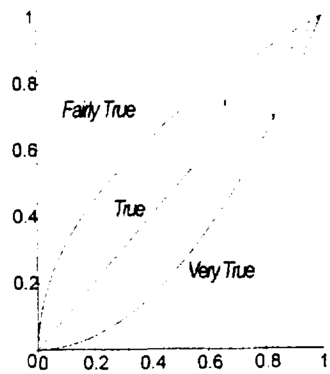


Figure 3: Some examples of linguistic truth values

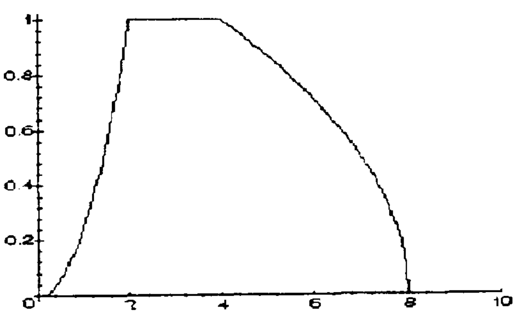


Figure 4: An example of generated membership functions

and

$$r(v) = \frac{V_d - v}{(V_d - V_c)}$$

When f is the square function and g is the square root function, the represented fuzzy number A is

$$A(v) = \begin{cases} 0 & \text{if } \; v < V_a \\ \left(\frac{v-V_a}{V_b-V_a}\right)^2 & \text{if } \; V_a \leq v < V_b \\ 1 & \text{if } \; V_b \leq v \leq V_c \\ \sqrt{\frac{V_d-v}{V_d-V_c}} & \text{if } \; V_c < v \leq V_d \\ 0 & \text{if } \; V_d < v \end{cases}$$

Figure 4 shows an example of these membership functions when $V_a = 0, V_b = 2, V_c = 4$ and $V_d = 8$.

3 Membership Function Generator: Lookup Table Approaches

Due to the parallel architecture of fuzzy rule block, very high performance is possible to achieve.. There is no limit in shapes of membership functions, and high execution speed. Gaussion, Sinusoidal, trapezoidal or any arbitrary shapes can easily be defined. To reduce size of the static RAM the number of membership functions can overlap [1]. This paper, however, describes architectures of a one RAM approach, and two RAM approach wit bout overlaps.

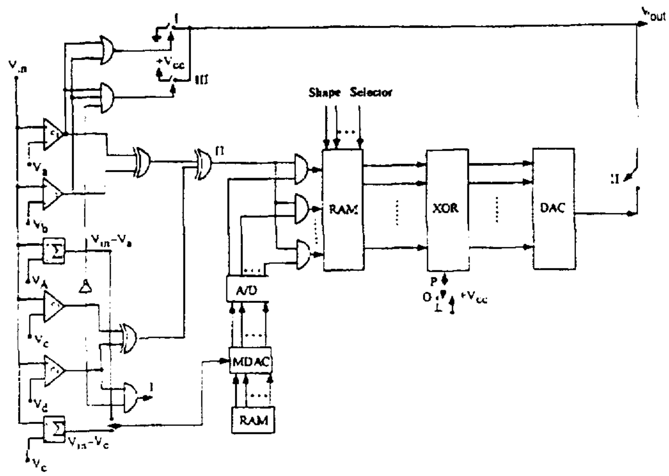


Figure 5: The block diagram of one RAM approach

3.1 Hardware Implemental ion with One RAM Approach

The block diagram of one RAM (or look up table) approach is shown in Fig. 5.

Signal V_{in}, to be fuzzified, is fed into simple comparator circuit Cl, C_2, C_3 and C_4. Signal $V_{in} - V_a$ and $V_{in} - V_c$ are multiplied by the values $\frac{1}{V_b - V_a}$ and $\frac{1}{V_d - V_c}$, which are stored in RAM, respectively. Multiplication between analog signal and "digital signal is done by a multiplying digit al to analog converter (MDAC). The digital output of analog to digital converter serves as address to RAM. As stated any type of function can be stored into RAM. Output of RAM is fed into XOR circuit. When polarity switch P is in position " O", values of RAM are restored at the output of the DAC. When polarity switch P is at position $+V_{cc}$, inverted values of RAM are restored at the output of the DAC. To save RAM memory space some additional switching circuitries are needed to generate low and high voltage.

3.2 Hardware Implementation with Two RAM Approach

In our two RAM approach, we use one RAM, RAM A, to store one half of the membership function which is referred as an "S-function," and the other RAM, RAM B, to store the other half of the membership function, which is referred as a "Z-function." An "S-function" can be represented by

$$L(v) = \begin{cases} 0 & if \quad v < V_a \\ l(v) & if \quad V_a \leq v < V_b \\ 1 & if \quad V_b < v \end{cases}$$

while a " Z-function" can represented by

$$R,(v) = \begin{cases} 1 & if \quad v < V_c \\ r(v) & if \quad V_c \leq v < V_d \\ 0 & if \quad V_d < v \end{cases}$$

where $l(v)$ and $r(v)$ are the left and the right functions of the membership function, respectively. Then, the original membership function is obtained by taking the minimum of the S-function and the Z-function. That is, the original membership function of a fuzzy set A *is* obtained by

$$A(v) = \min(L(v),\ R(v))$$

In the block diagram of the two RAM approach, shown in Fig. 6, the resulting membership function is obtained by taking minimum of the outputs from RAM A and RAM B. Note that the motivation of this approach was triggered by the analog approach given by Miki et al [3].

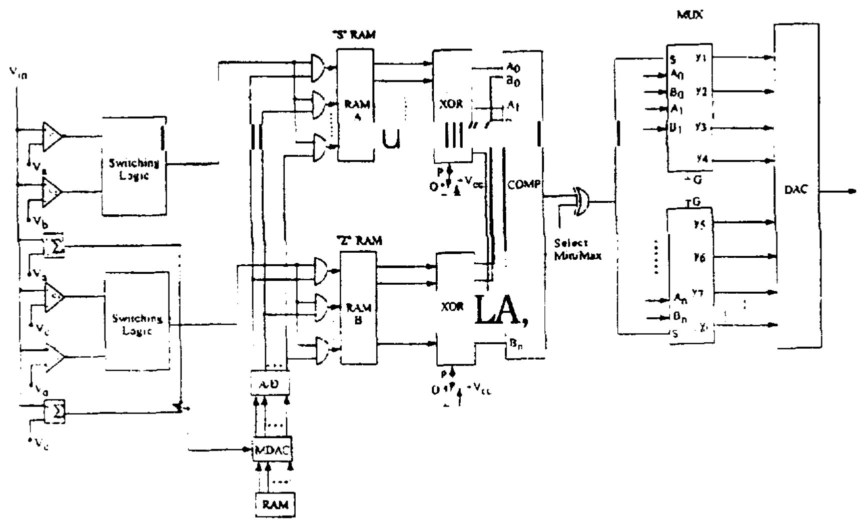


Figure 6: The block diagram. of two RAM approach

A horizontal shift of a membership function is achieved with programmable voltage thresholds, V_a, V_b, V_c, and V_d. Both RAM's are addressed at the same time. To obtained a membership function, a digital min/max circuit is realized. It consists of comparator, MUX and min/max selector. It is worth mentioning that two RAMs approach is complex but very flexible.

4 Membership Function Generator: Arithmetic Approaches

A key element in the design of a high speed fuzzy processor is the availability of a high speed fuzzifier to allow maximum processor (sometimes measured in number of fuzzy logic inference per second or FLIPS), A RAM approach is the answer to speed requirements.. Unfortunately, it is complex for VLSI realizat ion and power consuming as well. In this section, we discuss several efficient scheme for VLSI implement at ion using arithmetic approaches.

4.1 Piecewise Linear Membership Function: The Mult implication Approach

The proposed fuzzifier is characterized by a generalized trapezoidal membership function defined in terms of horizontal shifting parameters V_a, V_b, V_c, and V_d. Asymmetrical and symmetrical triangular S-shaped and Z-shaped curve are special cases of the generalized trapezoidal functions. Fig.7 presents a block diagram of a fast arithmetic membership function calculator. Preliminary results of this solution arc given in [4], its improved version is given in by Ramirez et al in this proceeding.

4.2 Piecewise Linear Membership Function: The Division Approach

Another approach is to use a division operation. Unlike the multiplication approach, the adjustable paramet ers in the division approach are only V_a, V_b, V_c and V_d. The block diagram of the division method is shown in Fig. 8. Detailed discussion including realization and simulation results of this approach is given in the paper " *Piecewise Linear Membership function Generator:A* Divider Approach' by R. Hart, G. Martinez, B. Yuan, D. Zrilic and J. Ramírez, which is also included in this proceeding. We should mention that this approach is simple from the realization point of view. But this approach is slower than the mult implication approach because division takes longer than multiplication.

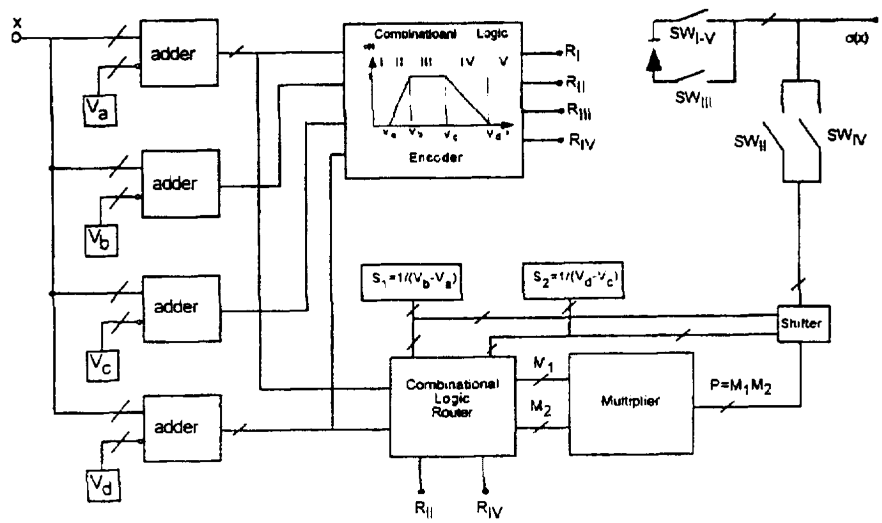


Figure 7 Scheme of a digital fuzzifier

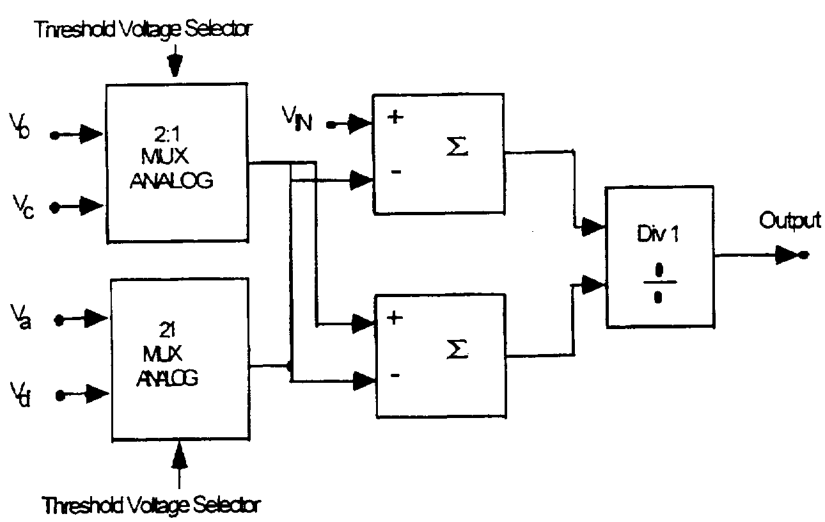


Figure 8: The block diagram of the divsion approach

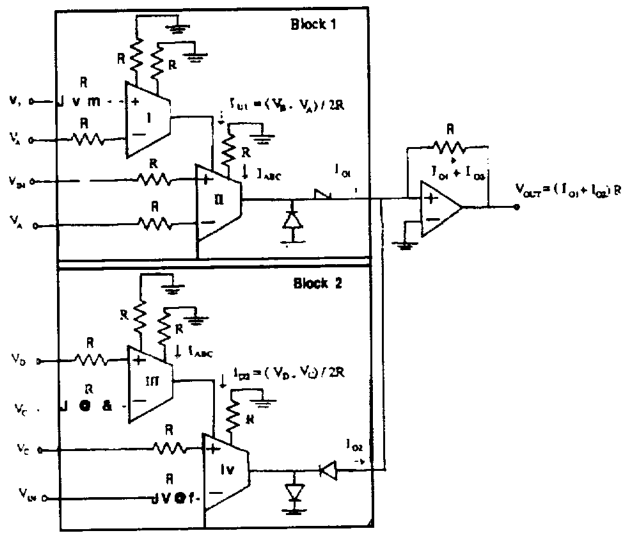


Fig-me 9: The circuit diagram of the current mode approach

4.3 Piecewise Linear Membership Function: The Current Mode Appreach

The basic building block of this type of membership function generator is operational transconductance amplifiers (OTA). An OTA is used as a voltage to current convert. The circuits diagram of thc current mode membership function generator is shown in Fig.9. It is easy to see that the circuit is simple and highly modular.. Experimental and simulation results are described in detail in the paper " *A new univers al analog fuzzifier based on operational transconduct ante* amplifiers" written by P. Saavedra, J. Ramírez-Angulo and J. Zrilic, which is also in this proceeding.

References

[1] M. Jacomet and R. Walti. A vlsi fuzzy processor with parallel rule execution. *5th IEEE Intern. Conference on Fuzzy Systems,* pages 554-5,58, 1996.

[2] G. J. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic: Theory and* Applications. Prentice Hall PTR, Upper Saddle River. NJ, 1995.

[3] T. Miki, H. Matsumoto, K. Ohto, and Y. Yamakawa. Silicon implimcntation for a novel high-speed fuzzy inference engine: Mega-flips analog fuzzy processor. *J. of Intelligent anf Fuzzy Systems,* 1(1):27–42, *1993.*

[4] J. Ramirez, N. Carneiro, D. Ma, and D. Zrilic. High speed digital vlsi fuzzifier. *The 39th Midwest Symposium on* Circuits *and Systems,* August 18-211996.

Page intentionally left blank

URC97145

W PHOTOPROTECTION IN TROPICAL MARINE ORGANISMS

Roy A. Armstrong
Tropical Center for Earth and Space Studies
University of Puerto Rico at Mayaguez

Introduction

Increasing levels of ultraviolet (UV) radiation reaching the earth's surface which results from stratospheric ozone depletions could have serious implications for terrestrial plants and for squat ic organisms within the euphotic zone. A documented 9% decline in ozone at mid-latitudes is considered to produce a 12% increase in harmful UV radiation (Kerr 1993). The biologically damaging effects of higher UV levels, particularly W-B (280-320 rim), could manifest earlier in the tropics because of the relative thinness of the earth's equatorial ozone layer. Tropical marine organisms are also living close to their upper tolerance levels of water temperature, However, despite the large potential effects on plants and animals, little is known about UV effects on tropical ecosystems. Long-term ecological studies are needed to quantify the effects of increased UV radiation on terrestrial and marine ecosystems and to produce reliable data for prediction.

Plants have developed several mechanisms to protect themselves from harmful UV radiation, one of which is the production of secondary leaf pigments that absorb W-B radiation (screening pigments). A higher concentration of screening pigments (e.g. flavonoids) in leaves may be interpreted as a natural response to increased *W* radiation. If higher concentrations of flavonoids filter out the excessive *W* radiation, no damage will occur, as suggested by Caldwell et al. (1989) and Tevini (1993). Failure to screen all W-B may result in deleterious effects on photosynthesis, plant genetic material, and plant and leaf morphology and growth. Eventually this will have an impact on ecosystem processes, structure, species composition, and productivity.

This paper describes an ongoing project that is assessing the responses of mangroves, seagrasses and corals to *W* radiation by studying pigment concentrations, biophysical parameters, and variations in spectral reflectance in the field and in W-reduction experiments. Preliminary results on the distribution of W-absorbing flavonoid compounds in red mangroves (*Rhizophora mangle*) and the seagrass *Thalassia testudinum*, *are* presented. This research also provides, for the first time, a permanent record of daily *W* irradiance measurements at a tropical location.

UV Effects on Plants

As suggested by Caldwell (198 1), Wellmann (1983), Beggs et al. (1986), and Braun (1991), anthocyanins and flavonoids have as one of their major functions the absorption of *W* radiation that might otherwise cause damage to the plant. Zeaxanthin, a widespread xanthophyll, is known to perform a protective role in plants by absorbing damaging *W* radiation (De Las Rivas et al. 1991). Increased *W* radiation affects photosynthesis rate and disrupts the chloroplast envelope (Barnes et al. 1987; Bornman et al. 1986; Sisson 1986; Campbell 1975; Caldwell et al. 1989; Tevini 1993). Damage seems to accumulate with duration of dose (Sisson 1986; Sullivan and Teramura 1992). Plant pathology is also augmented under increased UV-B radiation (Biggs and Webb 1986, Tevini 1993). Laboratory and field experiments in high latitudes have shown that increased W-B irradiance inhibits photosynthesis and increases accumulation of UV-absorbing pigments (Hardy et al. 1992). Pigment composition of red mangroves (Corredor, et al. 1995) closely follows the pattern found by De Las Rivas et al. (1989,1991) in several

species of deciduous trees from temperate environments further confirming the conservative nature of photosynthetic pigment suites in higher plants.

UV Effects on Marine Organisms

Both UV-A (320-400 nm) and UV-B (280-320 nm) are potentially important ecological factors in coral reefs (Jokiel and York 1982; Jokiel 1980). Reef building corals living near the equator tolerate higher UV levels than corals from high latitudes; shallow water corals also exhibit a higher UV tolerance than deeper corals (Maragos 1972). This adaptation results from variations in the concentration of a UV absorbing pigment, S-320, found in corals bearing endosymbiotic algae or zooxanthellae. Since the intensity of UV diminishes with increasing depth, less of this protective pigment is required at greater depths. The pigment S-320 is synthesized in response to UV light and not in response to other physical factors that also vary with depth (Jokiel and York 1982). Similar W-absorbing pigments have been extracted from marine algae (Sivalingham et al. 1974) and are probably common in other shallow water marine organisms.

Results

We have established a UV monitoring station in La Parguera, Puerto Rico using a Biospherical Instruments GUV-511 temperature-stabilized radiometer. This instrument has four bands centered at 308, 320, 340, and 380 nm in addition to PAR (400 -700 rim). Its purpose is to provide a high temporal resolution database of W irradiance. Data collection started in November 1996. Ozone data from the Total Ozone Mapping Spectrometer (TOMS) was acquired on-line from NASA's Climate Data System (NCDS) and from CD-ROM. Seasonal ozone variations in Puerto Rico show minimum values between November and the end of February (Figure 1). Maximum penetration of the shorter wavelengths of UV-B (<305 nm) is expected to occur at this time of the year. The highest surface irradiances of IN-A and the longer wavelengths of UV-B are expected to occur during the summer months, when lower solar zenith angles are present.

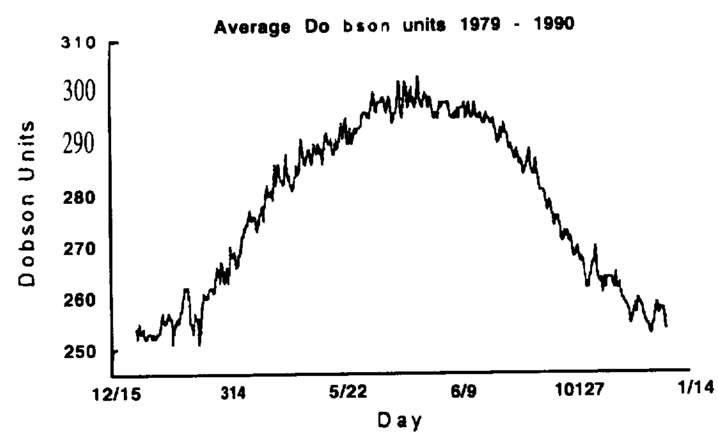


Figure 1: Seasonal ozone measurements using 12 year averaged data for Puerto Rico.

For this tropical location, the link between stratospheric ozone, solar zenith angle and surface UV spectral irradiance will be established after the first year of data collection by the UV monitoring station. For submerged plants and animals, such as seagrasses and corals, incident UV radiation also depends on water depth and the presence of dissolved and particulate organic matter in the water column. *In situ* UV measurements to a depth of 10 meters are being obtained with a Optronic Laboratories OL 754 spectroradiometer. This instrument offers an accuracy of ± 0.2 nm over the 200-800 nm spectral range (and ± 0.1 nm for the 280-320 nm range), is highly sensitive and has a large dynamic range, with user-selectable bandwidths of 1 to 10 nm.

Preliminary measurements of screening pigments in mangrove and seagrass leaves have been obtained by solvent extractions and spectrophotometric analysis. Figure 2 shows the absorption spectrum of W-absorbing pigments for upper-canopy (top) and shaded (bottom) red mangrove leaves. The sun-exposed upper-canopy leaves have higher amounts of photo protect ive pigments. Distinct peaks are present at about 280 and 330 nm.

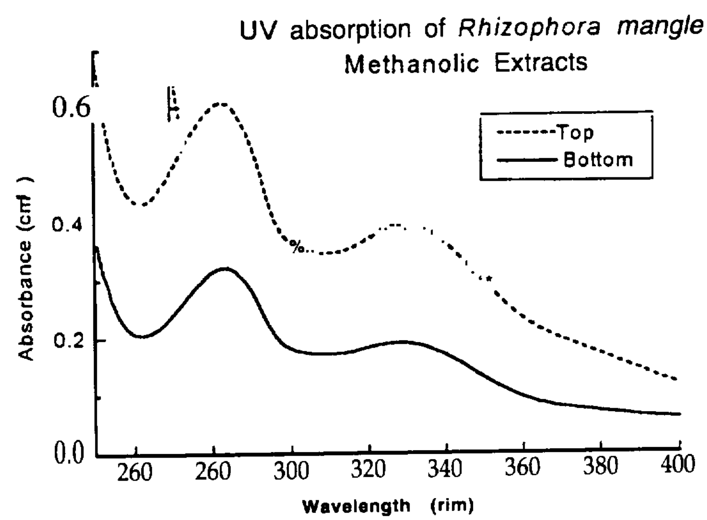


Figure 2: Absorption spectrum of photoprotective pigments in *R. mangle* leaves.

Pigment separation and identification in mangroves as well as seagrasses and corals is in process. This is being performed by a two-step isocratic HPLC/DAS following a modification of the procedure detailed by De Las Rivas et al. (1989, 1991). A C-18 reverse phase column is being used for pigment separation. For seagrasses and corals, the amount of photoprotective pigments is inversely proportional to water depth, due to the filtering effects of the water column. Figure 3 shows the absorption spectra of photoprotective pigments in the seagrass *T. testudinum* present at various depths and under full sun and shaded conditions. Shallower seagrasses have lesser amounts of UV-absorbing pigments. At the same depth, shaded seagrasses also have lesser amounts of protective pigments. A distinct absorption peak, most likely due to flavonoids, is present at about 330 nm.

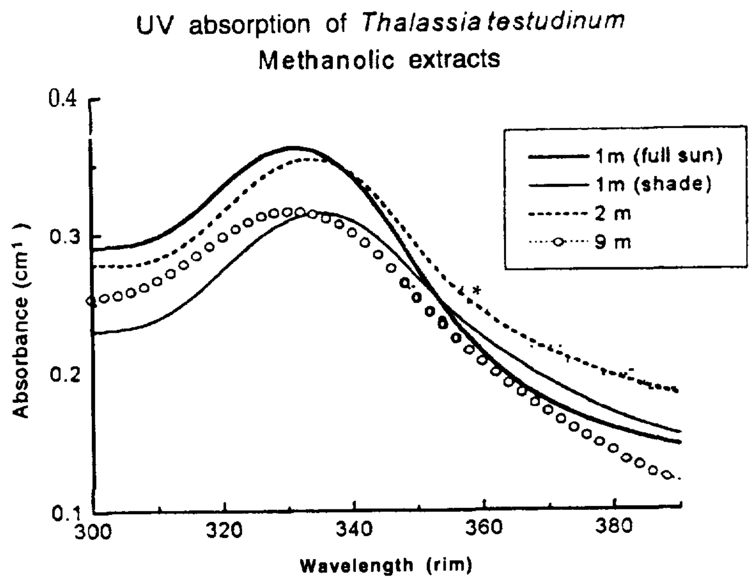


Figure 3: Absorption spectra of photoprotective pigments in *T. Testudinum.*

Conclusions

Preliminary results indicate that the total content of leaf photoprotective pigments, such as flavonoids in seagrasses and mangroves is irradiance dependent. Thus, for mangroves, lower canopy leaves exhibit lower flavonoid contents while upper canopy leaves exhibit higher contents. Similarly, seagrass flavonoid content is inversely correlated to depth with greater contents at shallow depths and vice-versa. Corals (or their symbiotic zooxanthellae) do not produce flavonoids but are known to produce a separate type of W-absorbing compound known as mycosporine-like amino acids or S-320's. S-320 content of corals is also irradiance dependent and increases dramatically upon bleaching; the stress-induced expulsion of zooxanthellae.

In the course of this study, we intend to document present levels of UV-photoprotective compounds in the major classes of sessile tropical marine organisms and to document variations in the levels of these pigments in response to natural and experimentally induced UV flux rates. This information will serve to assess the capacity of these organisms to modulate their response to variations in UV flux as well as to provide a baseline for assessment of the organismic response to expected increases in UV flux due to stratospheric ozone depletions.

Acknowledgments

This research was supported by grant NCCW-0088 from the National Aeronau [its and Space Administration. I thank J. Corredor, H. D'Antoni, X. Connelly, Y, Detrés, and F Muszynski for their help.

References

Barnes, P. W., Flint, S. D. and M. Caldwell. 1987. Photosynthesis damage and protective pigments in plants from aptitudinal Arctic/Alpine gradient exposed to supplemental UV-B radiation in the field. Arctic and Alpine Research 19 (1):21-27.

Beggs, C. J., U. Schneider-Ziebert and E. Wellmann. 1986, UV-B radiation and adaptive mechanisms in plants. in Worrest, R. and M. Caldwell, eds., Stratospheric ozone reduction, solar ultraviolet radiation and plant life. Springer Verlag.

Biggs, R. H. and P.G. Webb. 1986. Effects of enhanced ultraviolet-B radiation on yield, and disease incidence and severity for wheat underfield conditions. in Worrest, R. and M. Caldwell, eds., Stratospheric ozone reduction, solar ultraviolet radiation and plant life, Springer Verlag.

Bornman, J., R.F. Evert, R.J. Mierzwa. 1986. Fine structural effects of on UV-radiation on leaf tissue of Beta vulgaris. in Worrest, R. and M. Caldwell (editors) Stratospheric ozone reduction, solar ultraviolet radiation and plant life. NATO ASI Series. 374p.

Braun, J. 1991. The protective function of phenolic compounds of rye and oat seedlings against UV-B radiation and their biosynthetic regulation. in Tevini M. (cd.) Karlsr. Beitr. Entw. Oekophysiol. 9, 1-237.

Caldwell, M. M., A. H Teramura, and M. Tevini. 1989. The changing solar ultraviolet climate and the ecological consequences for higher plants. Tree 4 (12):363-366,

Caldwell, M.M. 1981. Plant response to solar ultraviolet radiation. in Lange, Nobel, Osmond and Ziegler (Eds.) Encyclopedia of Plant Physiology, Physiological Plant Ecology I, vol 12A. Springer Verlag.

Campbell, W.F. 1975. Ultraviolet-radiation-induced ultrastructural changes in mesophyll cells of soybean (Glycine max (L) Merr.). in Nachtwey, Caldwell and Biggs (editors) Impacts of Climatic Change on the Biosphere, Assessment Program, U.S. Department of Transportation Report No. DOT-TST-75-55, NTIS, Springfield, Virginia.

Corredor, J.E., E.J Klekowski, and J.M. Morell. 1995. Mangrove Genetics III. Pigment fingerprints of chlorophyll-deficient mutants. Int. J. Plant Sci. 156(1): 55-60.

De Las Rivas, J,, A Abadía and G. Abadía. 1989. A new reverse-phase HPLC method resolving ail major higher plant photosynthetic pigments. Plant Physiol. 91: 190-192.

De Las Rivas, J, J.C.G. Milicua and R. Gomez. 1991. Determination of carotenoid pigments in several tree leaves by reverse-phase high-performance liquid chromatography. J.

Chromat. 585: 168-172 .Harborne, J.E. 1992, Chromatography. 5th Edition. Chap 19, 363-393. Elsevier Publishing Co.

Hardy, T. J., M. Beherenfeld, H. Gucinski and A, Wones. 1992. Solar Ultraviolet Radiation in the South Pacific Ocean. AGU Ocean Sciences Meeting, 66 #041 C-7.

Jokiel, P.L. and R.H. York. 1982. Solar ultraviolet photobiology of the reef coral Pocillopora damicornis and symbiotic zooxanthellae. Bull, Mar. Sci. 32:301-315.

Jokiel, P.L. 1980. Solar ultraviolet radiation and coral reef epifauna. Science 207:1069-1071.

Kerr, R.A. 1993. Ozone takes a nose dive after the eruption of Mt. Pinatubo. Science 260, 490-491.

Lichtenhaler, H.K. 1987. Chlorophylls and carotenoids: Pigments of photosynthetic membranes. Meth. Ezymol. 148:350-382,

Mantoura, R.F.C. and C.A. Lewellyn. 1983. The rapid determination of algal chlorophyll and carotenoid pigments and their breakdown products in natural waters by reverse-phase high-performance liquid chromatography. Analyt. Chim. Acts 151:297-314.

Maragos, J.F. 1972. A study of the ecology of Hawaiian reef corals. Ph.D. thesis, Univ. of Hawaii, pp. 209.

Sisson, W.B. 1986. Effects of W-B radiation on photosynthesis. in Worrest, R and M. Caldwell, editors, Stratospheric ozone reduction, solar ultraviolet radiation and plant life, Springer Verlag.

Sivalingam P, T. Ikawa, Y. Yokohama, and K. Nisizawa. 1974. Distribution of a 334 nm absorbing substance in algae, with special regard of its possible physiological roles. Bet. Mar. 17:23-29.

Sullivan, J.H. and A.H. Teramura. 1992, The effects of ultraviolet-B radiation on loblolly pine. 2. Growth of field-grown seedlings. Trees 6, 115-120

Tevini, M. (Editor). 1993. UV-B Radiation and ozone depletion. Effects on humans, animals, plants, microorganisms, and materials. Lewis publishers.

Well mann, E. 1983. UV radiation in photomorphogenesis. in Shropshire Jr. W. and H. Mohr (eds.) Encyclopedia of Plant Physiology, new series, Photomorphogenesis, vol 1613. Springer Verlag.

Wright, S, W., S.W. Jeffrey, R.F.C. Mantoura, C.A. Llewellyn, T. Bjorland, D. Repeta and N. Welschmeyer. 1991. Improved HPLC method for the analysis of chlorophylls and carotenoids from marine phytoplankton. Mar. Ecol. Prog. Ser. 77:183-191

// /: /'

Application of Modular Modeling System To Predict Evaporation, Infiltration, Air Temperature, and Soil Moisture

By

Johnny Boggs, Latrica J. Birgan, Dr. Teferi Tsegaye, Dr. Tommy Coleman
[2]Dr. Vishwas Soman

Center for Hydrology, Soil Climatology, and Remote Sensing
Alabama A&M University
[2]Institute for Global Change Research and Education

ABSTRACT

Models are used for numerous application including hydrology. The Modular Modeling System (MMS) is one of the few that can simulate a hydrology process. MMS was tested and used to compare infiltration, soil moisture, daily temperature, and potential and actual evaporation for the Elinsboro sandy loam soil and the Mattapex silty loam soil in the Microwave Radiometer Experiment of Soil Moisture Sensing at Beltville Agriculture] Research Test Site in Maryland. An input file for each location was created to nut the model. Graphs were plotted, and it was observed that the model gave a good representation for evaporation for both plots. In comparing the two plots, it was noted that infiltration and soil moisture tend to peak around the same time, temperature peaks in July and August, and the peak evaporation was observed on September 15 and July 4 for the Elinboro Mattapex plot, respectively. MMS can be used successfully to predict hydrological processes as long as the proper input parameters are available.

INTRODUCTION

Mathematical modeling is an accepted scientific process providing a mechanism for comprehensively integrating basic processes (physical, biological, and chemical) and describing a system beyond what can be accomplished using subjective human judgments. As our understanding of the basic principles of basic processes deepens, it is possible to construct a model that better represents the natural system, and to use the models in an objective manner to guide both our future research efforts and the current measurement techniques. The soil properties and other hydrological processes are anexample of a natural system that has been modeled with different levels of resolution, and to which a modeling effort has been applied. Recently the Modular Modeling System (MMS) has been used to model the hydrology and related processes.

MMS is a comples system of computer software, written in FORTRAN of C language, to create models for various applications. MMS was developed in September 1989 with the establishment of a three-year agreement between the U. S. Geological Survey and the University of Colorado's Center for Advanced Decision Support for Water and Environmental Systems. MMS provides the framework needed to enhance development testing, and evaluation of physical-process algorithms and it facilitates the integration of user selected algorithms into operational physical-process models (MMS, User's Manual 1989).

Few hydrology models are currently available that simulate a hydrology process. The purpose of this project is to test MMS and simulate a hydrology process and to compare infiltration, soil moisture, daily temperature, and potential and actual evaporation for two bare plots at different locations consisting of Elinsbora sandy loam soil and Mattapex silty loam soil.

MATERIALS AND METHODS

The data used to run the MMS came for the Mocrowave Radiometer Experiment of Soil Moisture Sensing at Beltsville Agricultural Research Center Test Site from Maryland. Data for this project was collected everyday from June 19, 1981 to October 2, 1981 at approximately 9:30 a.m. MMS is designed to run on a Silicon Graphics computer using a UNIX system. To run and input data into the model, a basic knowledge of the commands are required. First, an imput file must be created for each bare plots. The input file requires header information with a list of variable names each followed by a number which represents the number of values for that variable in each row of the input file. Next, there should be a separator line which includes at least four pound symbols (####). After the separator line, the data lines are divided into fields, where each field is separated by a space. There are ten fields in the data line. The first six fields are assigned for time to include year, month, day, hour, minute, and second. The seventh, eight, ninth, and tenth fields are assigned for the input parameters rainfall (in.), minimum temperature ("C), maximum temperature (°C), and evaporation (in,).

After both input files are created, the model can be executed using one input file at a time. Once the proper procedure is followed to execute the model and input parameters are given to the model, the output data and plots will be generated. To make comparisons within and among plots the following four parameters were used 1) infiltration, 2) soil moisture, 3) average minimum, and maximum temperature, and 4) evaporation.

RESULTS AND DISCUSSION

By observing the graphs for the Elinboro Sandy loam soil (see Figures 1-4), several things can be noted. Infiltration has its highest peak on September 15, probably due to lack of runoff, nature of soil type and high rainfall. There is a steady decline in soil moisture from June until September where there is a sharp inclination that corresponds with the highest peak of infiltration. Temperature peaks in July and August, then it starts to decline. The highest peak of actual evaporation occurs on September 15. Although the potential evaporation that the model predicts is slightly overestimated, it is a good estimation of the actual evaporation. Overall as infiltration increases, soil moisture increases, but evaporation decreases. The decrease in evaporation could be attributed to a decrease in air temperature or due to a relatively high humidity.

Figure 1. Temporal representation of Infiltration for June 19, 1981 to October 2, 1981 for Elinsboro Sandy Loam Soil

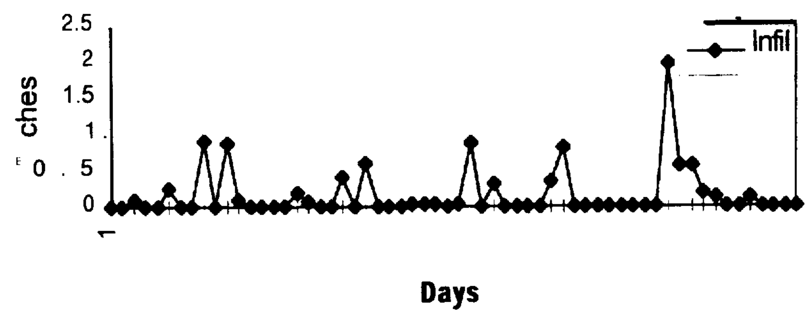


Figure 2. Temporal representation of soil moisture for June 19, 1981 to October 2, 1981 Elinsboro sandy Loam Soil

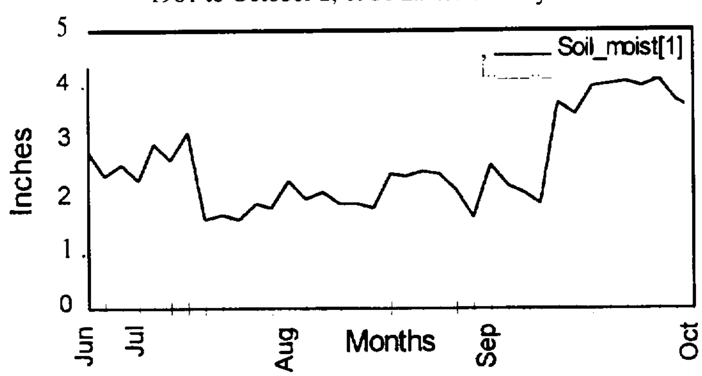


Figure 3. Temporal representation of Temperature (°C) for June 19, 1981 for Elinsboro Sandy Loam Soil

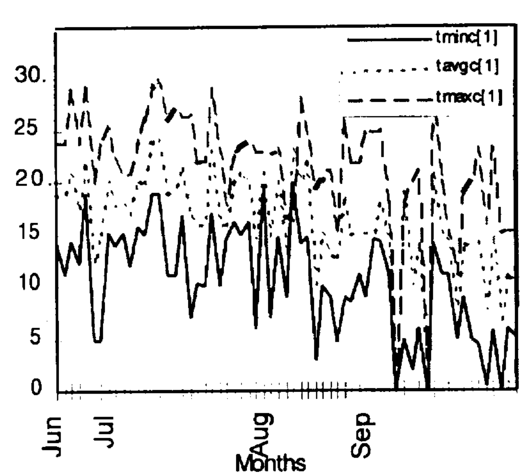
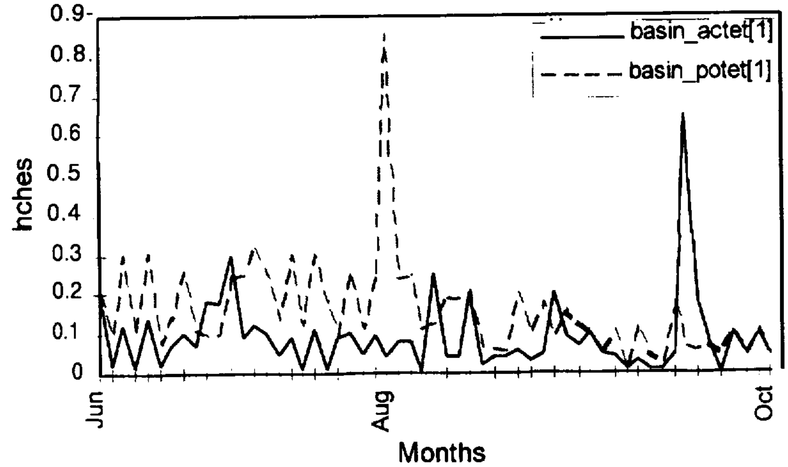


Figure 4 Actual and Potential Evaporation from June 19, 1981 October 2, 1981 for Bingoro Sandy Loam Soil



By studying the graphs for the Mattapex silty loam soil (see Figures 5-8), it can be shown that the highest infiltration peak values occur on July 3 and September 15, which could be attributed to lack of runoff and high rainfall. Soil moisture is also highest around the same time due to the preceding rainfall activity and less evaporation (data not shown), After July 3 there is a sharp decline in soil moisture until September 15, which is due to lack of rainfall. Temperature declines over time and has its highest peak on August 15, but its main peak time is during July and August. Actual evaporation peaks most around July 4. The model gives a good representation of evaporation.

Figure 5. Temporal representation of Infiltration for June 19, 1981 to October 2, 1981 for Mattapex Sandy Loam Soil

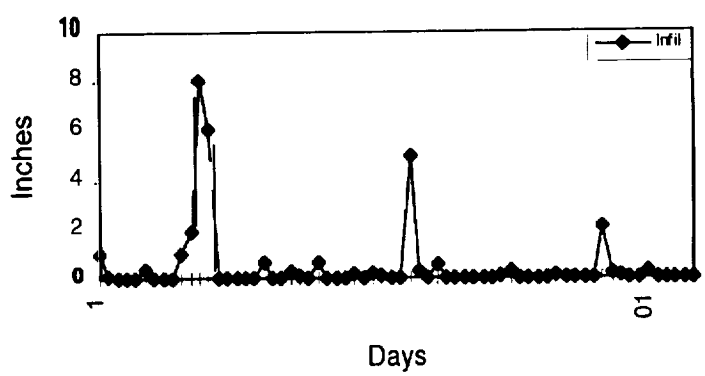


Figure 6. Temporal representation of Soil Moisture for June 19, 1981 to October 2, 1981 for Mattapex Silty Loam Soil

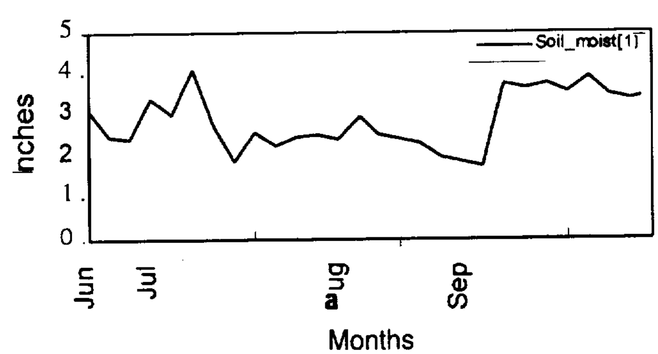


Figure 7. Temporal representation of Temperature (oC) for June 19, 1981 to October 2, 1981 for Mattapex Silty Loam Soil

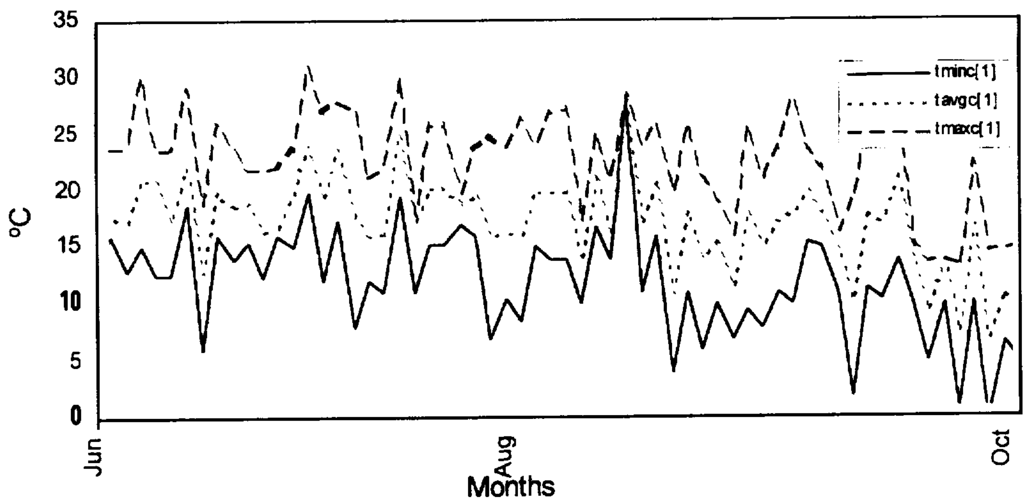
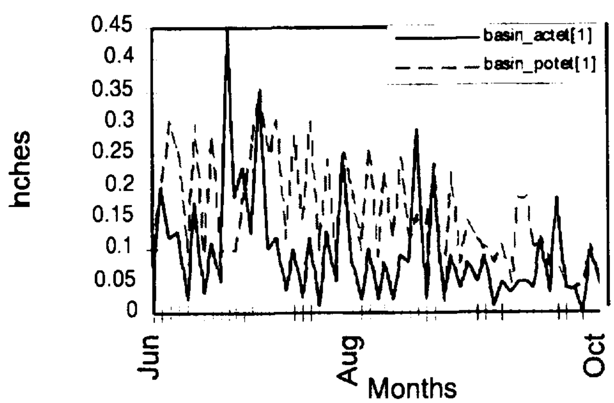


Figure 8. Actual and Potential Evaporation from June 19, 1981 to October 2, 1981 for Mattapex Silty Loam Soil



Comparing the infiltration, soil moisture, temperature, and evaporation of both experiment plots, many similarities exist among the plots, although the two plots are not in the same location. According to the graphs infiltration, soil moisture and temperature generally peak at the same time, but evaporation peaks at different times. This could happen because of the change in local variability of evaporation.

CONCLUSION

The MMS is currently running and an hydrology related process can be simulated. Comparing the output parameter of each plot, both plots behave similarly with very few dissimilarities. The highest peak of infiltration and soil moisture content for both plots was observed on September 15 as a result of the preceding rainfall acitivity on both locations. Temperature for both plots was at its highest peak in July and August, however, the peak evaporation for the two plots occurred at two different occasions i.e. for the Elinsboro plot it occurs on September 15, and for the Mattapex plot is occurs on July 4. Such models can be successfully used to predict soil surface evaporation and/or soil hydrology processes if the proper input data set is available.

ACKNOWLEDGEMENTS

Acknowledgment is extended to the Center for Hydrology, Soil Climatology, and Remote Sensing (HSCaRS) suppport staff. Contribution from HSCaRS, the Department of Plant, Soil and Animal Science, and the Alabama Agricultural Experiment Station, Alabama A&M University, Normal, AL 35762. Journal No, 343. This work was supported by Grant No. NCC W-0084 from the National Aeronautics and Space Administration (nASA), Washington, D.C. Any use of trade, product or firm names is for descriptive purposed only and does not imply endorsement by the U. S. Government.

REFERENCES

Modular Modeling System User's Manual: Volume One. University of Colorado at Boulder Center for Advanced Decision Support for Water and Environmental Systems 1989.

Wnag, J., T. Jackson, E. Engman, W. Gould, J. Fuchs, W. Glazer, P. O'Neil, T. Schmugge, and J. McMurtrey, 1984. Microwave Radiometer Experiment of Soil Mositure Sensing at BACR Test Site.

5146/46

USE OF ULTRASONIC TECHNOLOGY FOR SOIL MOISTURE MEASUREMENT

J. Choi, R.Metzl, M.D. Aggarwal, W. Belisle and T. Coleman
Center for Hydrology, Soil Climatology and Remote Sensing
Alabama A&M University, Normal, AL 35762

ABSTRACT

In an effort to improve existing soil moisture measurement techniques or find new techniques using physics principles, a new technique is presented in this paper using ultrasonic techniques. It has been found that ultrasonic velocity changes as the moisture content changes. Preliminary values of velocities are 676.1 m/sin dry soil and 356.8 m/s in 100 % moist soils. Intermediate values can be calibrated to give exact values for the moisture content in an unknown sample.

INTRODUCTION:

ultrasonic waves have been used extensively for material characterization and for sensing of material parameters[1,2]. Sound waves indeed are capable of providing useful information about the medium through which they travel. ultrasound is an extension of the audible sound and differs only in frequency range and not in principle. The waves produce small amplitude mechanical vibrations which are made to propagate through the sample under test. ultrasonic waves propagate through the material, the change their velocity and are attenuated depending on the material properties and are detecteda. The characteristics of the ultrasonic waves are modified as they travel through the material due to reflection scattering, and absorption. The detected signal can be displayed, processed and interpreted in terms of the properties of the material under investigation based upon its relation to the input wave. We feel that information about soil moisture can be obtained by measuring both the velocity and attenuation of the ultrasonic wave. The attenuation of an ultrasonic wave is associated with absorption and scattering of elastic waves by structural inhomogeneities. Scattering may be the governing attenuation mechanism in this medium. Because our interest is in measuring the moisture content in a variety of soil types, we have looked at the both the velocity difference and the attenuation.

The velocity of sound waves in water as determined in our experimental setup is 1558.6 m/s and in air it is 317.9 m/s measured at room temperature (21°C). It is estimated that the velocity of sound waves in dry soil and 100% moist soils differ by as much as 319.3 m/s and we will be able to interpolate an accurate scale for the determination of soil moisture.

PRINCIPLE OF ULTRASONIC TESTING

Consider an elastic medium as a network of atoms in a crystal lattice connected to each other by elastic forces. Suppose that a plane of the atoms at the surface of the medium is displaced by an external force following a harmonic function. All of these masses will undergo harmonic oscillation in the same phase. The elastic forces are

transmitted to the adjacent plane, then to the next adjacent plane, and so on. If the particles were rigidly coupled, the displacement and motion would be transmitted instantaneously. However, the elastic forces introduce a time delay that increases with distance. The above described phenomenon is an elastic wave, which is called an ultrasonic wave if its frequency is higher than 20 kHz. Two modes of propagation are possible in an infinite medium:

(1) longitudinal waves or pressure waves by which particles are displaced along the direction of propagation,

(2) shear waves or transverse waves by which the particles are displaced perpendicular to the direction of propagation.

In longitudinal waves, the sound velocity c_l is determined by the density p, the modulus of the elasticity E, and the Poisson ratio σ in an infinite medium[3,4]:

$$c_l = \sqrt{\frac{E}{\rho}\frac{1-\sigma}{(1+\sigma)(1-2\sigma)}}, \tag{1}$$

$$c_l = \sqrt{\frac{v}{\rho}}, \tag{2}$$

where $v = \lambda' + 2\mu$, λ' and μ are Lamé constants.

Considering plane waves, attenuation A after traveling a distance r with incident amplitude A_0 is;

$$A = A_0 e^{-\alpha r} \tag{3}$$

The attenuation coefficient a is given by

$$\alpha = \frac{1}{r}\ln\frac{A_0}{A} \tag{4}$$

For ultrasonic waves in a homogeneous medium, part of the wave energy is absorbed and turns into heat. In a heterogeneous medium, additional losses take place in the form of scattering. The attenuation coefficient is given by

$$\alpha = \alpha_a + \alpha_s \tag{5}$$

where α_a ***is*** the absorption coefficient and α_s is the scattering coefficient.

In solids, absorption is dominated by elastic hysteresis and internal friction. These mechanisms have frequency dependence that is normally a linear function in solids. Scatter is due to inhomogeneities that depends on the wavelength and a characteristic impedance different from that of the surrounding materials, The scattering coefficient

varies with the wavelength-to-diameter ratio of the inhomogeneities. In summary, the value of the velocities will be different for different soils

Experimental Setup

A schematic diagram of the setup designed and fabricated in the laboratory is shown in Fig. 1. It consists of a Panametrics pulser/receiver unit combining an impulse type pulser with a receiving system and the sample holder for the soil sample under investigation as well as an oscilloscope to accurately measure the travel time for the ultrasonic waves.

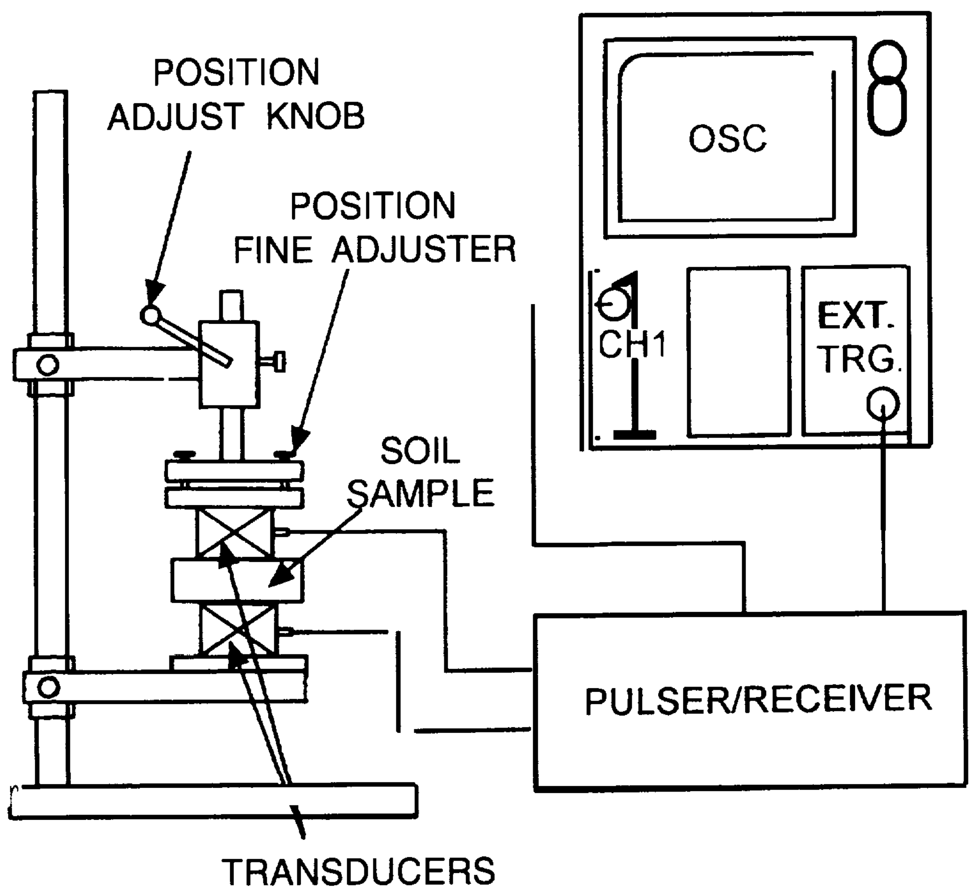


Fig. 1. "Experimental setup for measuring soil moisture using ultrasonic transducers

Two transducers are attached to a vertical support and the sample can be sandwiched between them. Coupling of the sample to the transducer has been found to be critical to control conduction of the sound between tranducer and sample. Normally coupling to the sample is accomplished by using a light weight oil, glycerin or other suitable couplant. In the present case, we have used dry contact coupling which introduces some uncertainties in the measurement. The analog recording system makes it

difficult to determine the travel time accurately, The angle of incidence is kept perpendicular to the surface of the sample and the time of travel is measured. A transmitting 500 kHz transducer is connected to the pulser to apply electric pulses. The pulses are then converted to the ultrasonic signal in the transducer. The receiving transducer is connected to the receiver to amplify the received signal, This amplified signal is displayed on the oscilloscope screen.

For a typical measurement, the following procedure is followed:

(1) Two transducers are placed face to face (zero distance) and a reference peak is recorded,

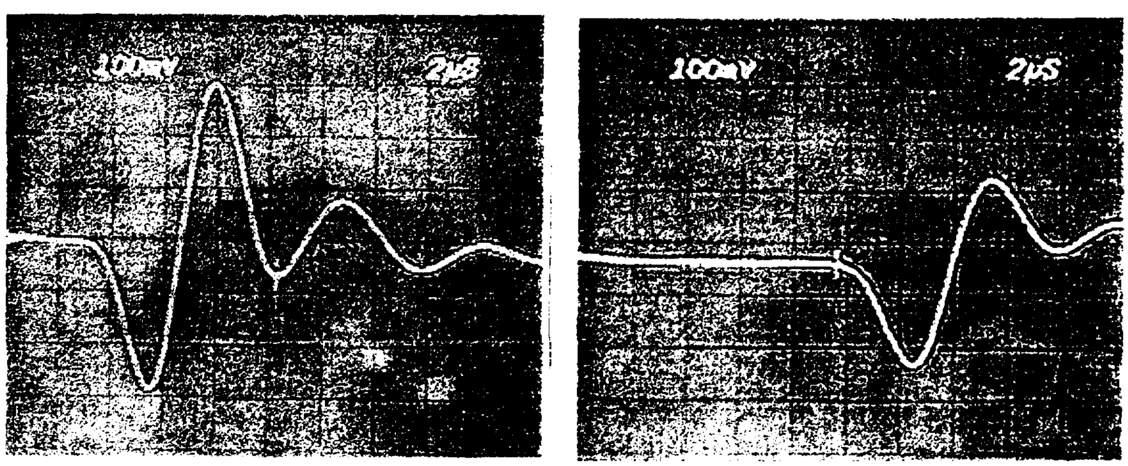
(2) A sample is placed between the transducers and one of the transducers is adjusted to get a maximum amplitude and symmetrical sinusoidal waves without distortions.

(3) A propagation time is calculated by reading the time between the reference peak and the peak from the received signal.

(4) Ultrasonic velocity is calculated using the propagation time and the length of the soil sample under investigation.

Results and Discussion

Preliminary results for the measurement of ultrasound velocities in Decatur soil are given in Table 1, and are calculated by measuring the time interval from the reference point on the oscilloscope screen, Fig. 2 below shows the oscilloscope traces for a) dry soil and b) wet soil.



(a) Dry Soil (b) Wet Soil

Fig. 2 Cathode Ray Oscilloscope traces to determine the travel time through the samples

Typical values of ultrasonic velocities in soils and other media are listed in Table 1.

Table 1. Typical measured ultrasonic velocities in soils and other media

Specimen	Sample Thickness (mm)	Time (μS)	Longitudinal Velocity (m/s)
Air	1.9	6.0	317.9
Water	11.2	7.2	1558.6
Decatur Soil (Dry) ,	3.3	4.85	676.1
Decatur Soil (Wet)	4.78	13.4	356.8

Different values of ultrasonic velocity have been found for different packing for the same composition and moisture content. Attenuation of the ultrasonic signal is a function of the particle size and hence may be used as a possible method for particle size determination. Results show a differential in ultrasonic velocity of 319.3 m/s between dry and 100% moist soil. These results are planned to be used for making a moisture sensor. The relationship between ultrasonic velocity and moisture content as a function of packing also needs to be studied, More research is needed to bring this technology to fruition and work is underway.

Acknowledgments

Acknowledgment is extended to the Center for Hydrology, Soil Climatology and Remote Sensing (HSCaRS) support staff. Contribution from HSCaRS, the Department of Soil, Plant and Animal Science, and the Alabama Agricultural Experiment Station, Alabama A&M University, Normal, AL 35762. Journal No. 345. The work is supported by Grant No. NCC-0084 from the National Aeronautics and Space Administration (NASA), Washington, D.C. Any use of trade, product or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References:

1. T.E. Matikas, R.L. Crane, Mat. Res. Sot. Bull. 21, 18 (1996)
2. S.I. Rokhlin and T.E. Matikas, Mat. Res. Bull. 21,22 (1996)
3. E.P. Papadakis, in Physical Acoustics, Principles and Methods edited by W.P. Mason and R.N. Thurston Vol XII Academic Press, New York 1976 p. 277.
4. M.C. Bhardwaj, Advanced Ceramic Materials 1,311 (1986)

Page intentionally left blank

URC97148

Optical Characterization and 2,525 μm Lasing of Cr^{2+}: $Cd_{0.85}Mn_{0.15}Te$

5/47/ 7?

V. R. Davis, X. Wu, and U. Hömmerich
Research Center for Optical Physics,
Department of Physics
Hampton University
Hampton, VA 23668

S. B. Trivedi, K. Grasza, and Z. Yu
Brimrose Corporation of America
5020 Campbell Blvd.
Baltimore, MD 21236

1. Introduction

Transition metal doped solids are of significant current interest for the development of tunable solid-state lasers for the near and mid-infrared (1-4 pm) spectral region [1,2]. Applications of these lasers include basic research in atomic, molecular, and solid-state physics, optical communication, medicine, and environmental studies of the atmosphere.

In transition metal based laser materials, absorption and emission of light arises from electronic transitions between crystal field split energy levels of 3d transition metal ions. The optical spectra generally exhibit broad bands due to the strong interaction between dopant and host (electron-phonon coupling) [3]. Broad emission bands offer the prospect of tunable laser activity over a wide wavelength range, e.g. the tuning range of Ti: Sapphire extends from 700-1100 run [1,2]. The only current transition metal laser operating in the mid-infrared wavelength region (1.8-2.4 μm) is Co^{2+}: MgF_2, but its performance is severely limited due to strong nonradiative decay at room temperature. Based on lifetime data, the quantum efficiency is estimated to be less than 3°/0 [1,2]. In general, the probability for non-radiative decay via multi-phonon relaxation increases with decreasing energy gap between ground and excited state. Therefore, efficient transition metal lasers beyond -1.6 μm are rare [1,2].

Recently, tunable laser activity around 2.3 μm was observed from Cr doped ZnS and ZnSe [4,5]. The new lasing center in these materials was identified as Cr^{2+} occupying the tetrahedral Zn site. Tetrahedrally coordinated optical centers are rather unusual among transition metal lasers [1,2]. Their potential usefulness, however, has been demonstrated by the recent development of near infrared laser materials such as Cr: forsterite and Cr: YAG, which are based on tetrahedrally coordinated Cr^{4+} ions [6,7]. According to the Laporte selection rule, electric-dipole transition within the optically active 3d-electron shells are parity forbidden [3]. However, a static acentric electric crystal field or the coupling of asymmetric phonons can force electric-dipole transitions by the admixture of wavefunctions with opposite parity. Tetrahedral sites lack inversion symmetry which provides the odd-parity field necessary to relax the parity selection rule. Therefore, high absorption and emission cross sections are observed. An enhanced radiative emission rate is also expected to reduce the detrimental effect of non-radiative decay [4,5]. Motivated by the initial results on Cr doped ZnS and ZnSe, we have started a comprehensive effort to study Cr^{2+} doped II-VI semiconductors for solid-state laser applications. In this paper we present the optical properties and the demonstration of mid-infrared lasing from Cr doped $Cd_{0.85}Mn_{0.15}Te$.

2. Experimental details

The undoped and Cr doped $Cd_{0.85}Mn_{0.15}Te$ samples were prepared at Brimrose Corporation using a modified Bridgman growth method. The nominal Cr concentration was $5x10^{19}$ atoms/cm^3. Absorption measurements were carried out using a Cary 5E spectrophotometer. Luminescence was excited with the 1.9 μm output of a Q-switched Nd: YAG pumped Optic-al Parametric Oscillator. The infrared emission was dispersed with a 1m monochromator and detected with a liquid nitrogen cooled InSb detector. The emission signal was processed with a boxcar averager and personal computer. All spectra were corrected for the spectral response of the luminescence setup. Lifetime measurements were performed by directly monitoring the InSb signal on an averaging digitizing oscilloscope. Cooling was achieved by mounting the sample on the cold finger of a closed cycle helium refrigerator. Laser experiments were performed using the pulsed (10Hz) 1.9 μm output of a H_2-Raman shifted Nd: YAG laser.

3. Results and Discussion

Absorption:

The unpolarized absorption spectra of undoped (upper trace) and Cr doped (lower trace) $Cd_{0.85}Mn_{0.15}Te$ are shown in Figure 1. The incorporation of Cr into the $Cd_{0.85}Mn_{0.15}Te$ lattice results in a strong absorption band centered at around 1900 nm with a width of -400 nm (FWHM). This absorption feature is similar to that measured for other Cr doped II-VI's semiconductors [8] and is attributed to the only spin-allowed transition ($^5T_2 \rightarrow {}^5E$) of Cr^{2+} ions ($3d^4$ electronic configuration) in a tetrahedral telluride coordination. Using the nominal Cr concentration of $5\times10^{19}/cm^3$ the peak absorption cross section was calculated to be 4.4×10^{-19} cm^2. This value has to be considered as a lower limit because the actual concentration of Cr^{2+} ions in the sample is unknown.

Figure 1 also shows that the bandedge of Cr doped $Cd_{0.85}Mn_{0.15}Te$ exhibits a strong absorption shoulder extending nearly down to 1000 nm which is not observed in the undoped $Cd_{0.85}Mn_{0.15}Te$ sample. Since all transitions from the $^5T_2$ ground state of Cr^{2+} to higher excited states are spin-forbidden, it is unlikely that the 1000 nm absorption shoulder arises from an intra 3 d transition. Further work on the identification of the Cr induced near bandedge absorption is in progress.

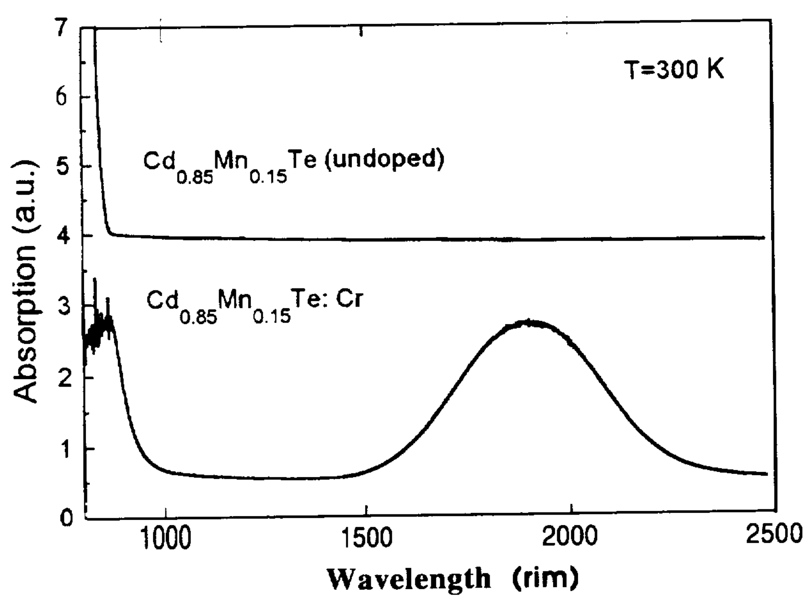


Figure 1: Unpolarized absorption spectra of undoped (upper trace) and Cr doped (lower trace) $Cd_{0.85}Mn_{0.15}Te$. The strong absorption band centered at 1900 nm is attributed to tetrahedrally coordinated Cr^{2-} ions.

Emission:

The room temperature luminescence spectrum of Cr: $Cd_{0.85}Mn_{0.15}Te$ excited at 1900 nm shows a strong band centered at 2250 nm with a FWHM of 450 nm and a lifetime of 1.4 µs. When cooling the sample to 15 K, the emission band narrows and consists of a main peak at 2200 nm and a shoulder at 2400 nm. The double peak in the low temperature spectrum indicates the

derivation of the local Cr^{2} environment from perfect tetrahedral symmetry [3]. At 15 K the lifetime increases to 3.7 µs. Under the assumption that the radiative decay rate is constant with temperature, any change in he measured lifetime is interpreted as being due to the onset of non-radiative decay [3,9], This yields a promising high quantum efficiency of -38% at room temperature.

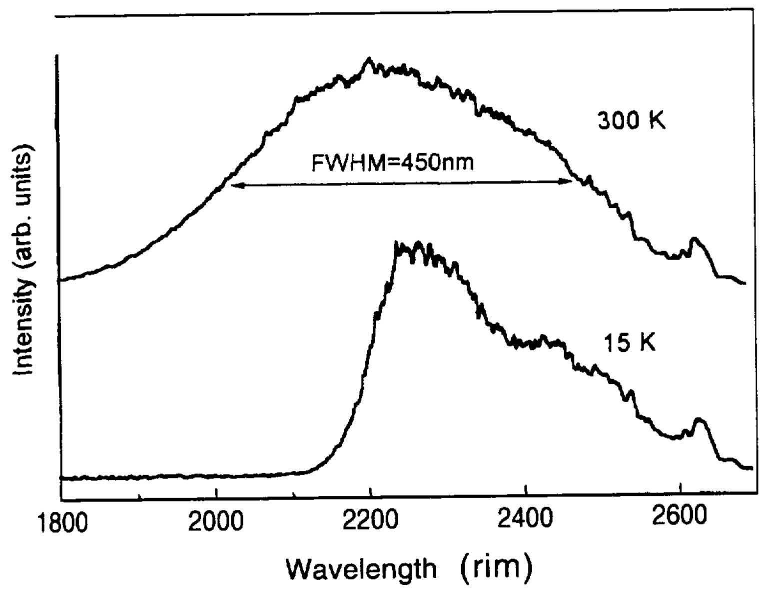


Figure 2: Mid-infrared luminescence spectra of Cr^{2+}: CdMnTe at low and room temperature. The luminescence was excited with the 1900 nm output of a Q-switched Nd: YAG pumped Optical Parametric Oscillator.

Using the estimated quantum efficiency and lifetime τ, the room temperature emission cross section σ_{em} was calculated according to McCumber's theory [10,11]:

$$\sigma_{em} = \eta \cdot \sqrt{\frac{\ln 2}{\pi}} \frac{1}{4\pi c n^2 \tau} \frac{\lambda^4}{\Delta\lambda}$$

where n is the refractive index, c is the speed of light, λ is the emission wavelength, and $\Delta\lambda$ is the full width half maximum (FWHM) of the emission band. It is implied in McCumber's theory that the emission spectrum has a nearly Gaussian band-shape. For the case of Cr: $Cd_{0.85}Mn_{0.15}Te$, we obtained an emission cross section of 2.7×10^{-18} cm^2, using η=38%, λ=2250 nm, $\Delta\lambda$=450 nm, τ=1.4µs, and n=2.7 (CdTe). The emission cross section of Cr^{2+}: CdMnTe is significantly larger than that of the commercial laser material Ti: Sapphire [1,2].

Laser experiments

Room temperature laser operation of Cr^{2+}: CdMnTe was demonstrated with a 3-4 mm thick sample placed in a cavity consisting of a flat high reflector (R>99% @ 2350 nm) and a curved output coupler (R=95% @ 2350nm). For the excitation of Cr^{2+} ions the 1907 nm output from a H_2 Raman-Cell pumped by a Nd: YAG laser was used (Figure 3). Laser activity centered at 2525 nm was achieved with a slope efficiency of 4. **5°/0** (Figure 4). More detailed studies on the laser performance of Cr^{2+}: $Cd_{0.85}Mn_{0.15}Te$ are currently in progress and will be published in a forthcoming paper [12].

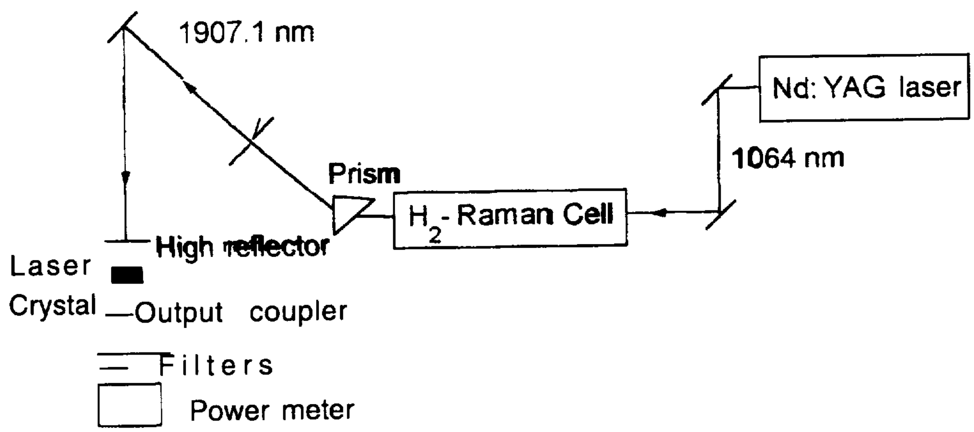


Figure 3: Experimental setup for testing the laser activity of Cr^{2+}: CdMnTe.

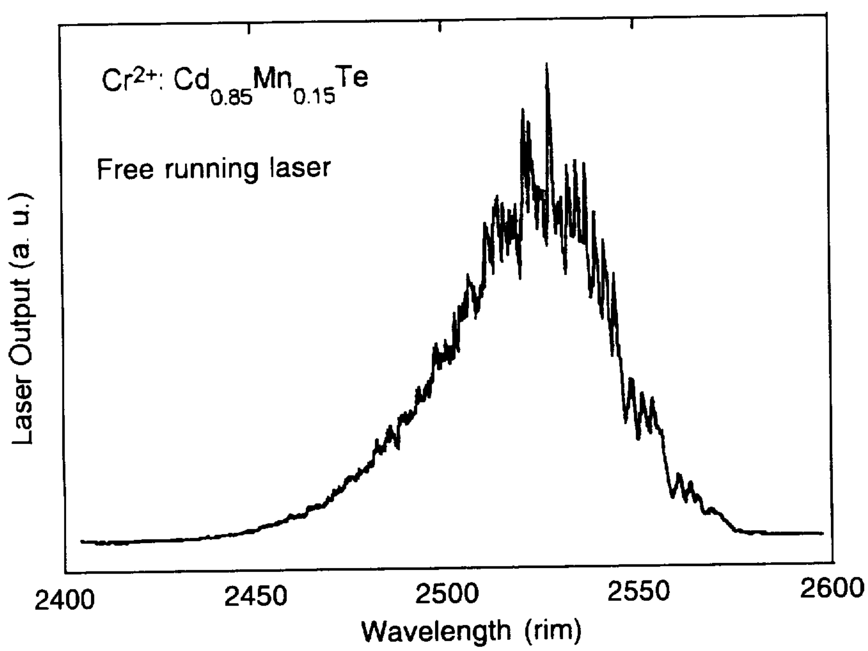


Figure 4: Spectrum of the free running Cr^{2-}: $Cd_{0.85}Mn_{0.15}Te$ laser at room temperature.

4. Conclusion

The optical properties of Cr doped $Cd_{0.85}Mn_{0.15}Te$ were presented. This system shows a broad absorption band centered at 1900 nm which is attributed to an intra 3d transition of tetrahedrally coordinated Cr^{2+} ions. Direct excitation into this absorption band resulted in a strong mid-infrared luminescence centered at 2250 nm. Based on lifetime measurements the quantum efficiency was estimated to be 3 **8°/0.** Initial laser experiments have been carried out and room temperature laser activity at 2525 nm with a FWHM of -50 nm was demonstrated. Under less than optimal conditions the slope efficiency was measured to be 4. 5%.

Acknowledgment

The authors from Hampton University would like to acknowledge the financial support of NASA through grant No. NAGW-2929.

References

[1] P.F. Moulton, Proceedings of the IEEE, Vol. **80**, No. 3 (1992) 348.
[2] F. J. Duarte, *Tunable Laser Handbook*, Academic Press, San Diego, 1995
[3] B. Henderson and G. F. Imbusch, *Optical Spectroscopy of inorganic Solids*, Clarendon Press, Oxford, 1989.
[4] L. D. DeLoach, R. H. Page, G.D. Wilke, S. A. Payne, and W. F. Krupke, OSA Proceedings on Advanced Solid-State Lasers, Bruce H. T. Chai and Stephen A. Payne (eds.), 1995, Vol. 24, pp. 127-131.
[5] L. D. DeLoach, R. H. Page, G. D. Wilke, S. A. Payne, and W.F. Krupke, IEEE J. Quantum Electron. Vol. 32, No.6, 1996, 885.
[6] V. Petricevic, S. K. Gayen, and R. R. Alfano, Appl. Phys. Lett. 53 (1988) 2590.
[7] S. Kück, J. Koetke, K. Petermann, U, Pohlmann, and G. Huber, OSA Proceedings on Advanced Solid-State Lasers, Albert A. Pinto and Tso Yee Fan, eds., 1993, Vol. **15**, pp. 334-338.
[8] J. T. Vallin, G. A. Slack, S. Roberts, and A. E. Hughes, Phys. Rev. B, Vol. 2, pp. 4313-4333, 1970.
[9] B. Di Bartolo, *Optical Interactions in Solids*, Wiley, *New* York, 1968.
[10] D. E. McCumber, Phys. Rev. 134, A299 (1964).
[11] S. Kück, K. Petermann, U. Pohlmann, and G. Huber, Phys. Rev. B> 51,1995, 17323.
[12] U. Hömmerich et al., to be published.

Page intentionally left blank

URC97149

CALMODULIN-DEPENDENT PROTEIN KINASE MEDIATES HYPERGRAVITY-INDUCED CHANGES IN F-ACTIN EXPRESSION BY ENDOTHELIAL CELLS,

Felisha D. Love[§], Caroline Melhado[§], Francis Bosah[‡], Sandra A. Harris-Hooker[‡] and Gary L. Sanford[§]

Departments of Biochemistry[§] and Medicine[‡]

Morehouse School of Medicine, Atlanta, GA 30310

INTRODUCTION: A number of basic cellular functions, e.g., electrolyte concentration cell growth rate, glucose utilization, bone formation, response to growth stimulation and exocytosis are modified by microgravity or during spaceflight (1-4). Studies with intact animal during spaceflights have found lipid accumulations within the lumen of the vasculature and degeneration of the vascular wall (5). Capillary alterations with extensive endothelial invaginations were also seen (6). Hemodynamic studies have shown that there is a redistribution of blood from the lower extremities to the upper part of the body; this will alter vascular permeability, resulting in leakage into surrounding tissues (7). These studies indicate that changes in gravity will affect a number of physiological systems, including the vasculature. However, few studies have addressed the effect of microgravity on vascular cell function and metabolism. A major problem with ground based studies is that achieving a true microgravity environment for prolonged period is not possible. On the other hand, increasing gravity (i.e., hypergravity) is easily achieved. Several researchers have shown that hypergravity will increase the proliferation of several different cell limes (e.g., chick embryo fibroblasts) while decreasing cell motility (8) and slowing liver regeneration following partial hepatectomy (9). These studies suggest that hypergravity will alter the behavior of most cells. Several investigators have shown that hypergravity affects the expression of the early response genes (c-fos and c-myc) and the activation of several protein kinases (PK's) in cells (1 0,11). In this study we investigated whether hypergravity alters the expression of f-actin by aortic endothelial cells, and the possible role of protein kinases (calmodulin(II)-dependent and PKA) as mediators of these effects.

EXPERIMENTAL METHODS: BAECs were obtained from NIGMS/Coriell Cell Repository. Cells were maintained in Dulbecco's Modified Eagle's Medium supplemented with 10% fetal bovine serum and 1X antibiotics. We assessed F-actin expression in BAECs subjected to centrifugation in the presence or absence of protein kinase inhibitors (PKIs).

F-actin Expression When Subjected to Hypergravity Subconfluent cultures were subjected to hypergravity (centrifuged at 6 and 12 G) for 24,48 and 72 hr. The cultures were rinsed with phosphate buffered saline (PBS) and then fixed with 1%glutaraldehyde/0.1% Triton X-1 00/2.5% Formalin in PBS for 24 hours. Following fixation, the cells were stained with FITC-phalloidin and examined using phase and fluorescence microscopy. Controls were treated similarly except they were not subjected to centrifugation.

Effectof PKIs on F-actin Expression Under Hypergravity: The PKIs, KT5926 and KT5720, were added, singly, to the medium of subconfluent cultures (2uL PKI/mL) and centrifuged at 6 and 12 G for 24,48, and 72 hr. Cultures were stained and examined for f-actin expression as previously outlined. Controls were treated similarly except they were not subjected to centrifugation.

RESULTS AND DISCUSSION: Figure 1 shows subconfluent BAECs that were subjected to hypergravity (12G). The cells generally showed increased F-actin expression throughout the cell as a function of time under hypergravity. Optimal fluorescence was observed after 48 hr with a slight decline seen after 72 hr. Cells subjected to 6G of hypergravity gave similar results. The controls, however, generally showed a decreased F-actin expression with increasing time of culture, with no fluorescence observed after 72 hr. The fluorescence observed for control cells

was localized to the nuclear region instead of throughout the cell as observed for hypergravity treated cells. There were no other morphological differences seen for either hypergravity-treated or control cells when examined by phase contrast microscopy. In this study, we only examine changes in f-actin. These findings suggest that hypergravity has a direct effect on f-actin levels in endothelial cells but could also affect other forms of actin. The latter possibility will be investigated in future studies.

Subconfluent BAECs subjected to hypergravity in the presence of the PKI - KT5926, which is a specific inhibitor of calmodulin(II)-dependent protein kinase, exhibited increased f-actin expression with increasing time under 6G is shown in Figure 2. Similar results were observed for cells subjected to 12G hypergravity. The fluorescence, observed under either 6 or 12G, was especially noted at cell-to-cell junctions as well as for the nuclear regions. The controls for PKI-KT5926 showed a slight increase in f-actin expression with increasing time under 6G. The protein kinase inhibitor KT5720 did not affect f-actin expression by BAEC under any of the conditions studied. Again, no other morphological changes were observed when examined by phase contrast microscopy. These results indicate that, of the two protein kinases examined, only the calmodulin(II)-dependent protein kinase has a mediator roles for cells subjected to hypergravity. Since calmodulin(II)-dependent protein kinase is activated normally by either increased intracellular Ca^{+2} concentration or inositol-3-phosphate, these may also mediate hypergravity-induced cellular changes. However, with the complexity for signal transduction mechanisms, it is not possible with the present study, to venture farther in possible signaling events.

ACKNOWLEDGEMENTS: This study was supported by grants NASA NAG9-644 and NCCW-008.

REFERENCES

1. Rijken PJ, et al. (1991). Aviat Space Environ Med 62:32-36.
2. Kumei Y, et al. (1989). J Cell Sci 93:221-226.
3. Cogoli A, Tschogg A and Fuchs-Bislin P (1984). Science 225:228-230.
4. Gruener R and Hoeger G (1990). Am. J Physiol 258: C489-C494.
5. Doty S, E Holton, G Durnova and A Kaplansky. (1990) FASEB J 4:16.
6. Philpott D, I Popova, K Kate, et al. (1990) FASEB J 4:73.
7. Nixon J, R Murray, C Byrant, et al. (1979). J Appl Physiol 46:541.
8. Tschopp A and Cogoli A, Experiential 39, 1323-1329, 1983
9. Kropacova K, et al., The Physiologists 31, S75-S76, 1988
10. Nose K and Shibanuma M, Exp. Cell Res. 211:168-70, 1994
11. DeGroot KP, et al., Exp. Cell Res. 197:87-90, 1991.

A B C D E F

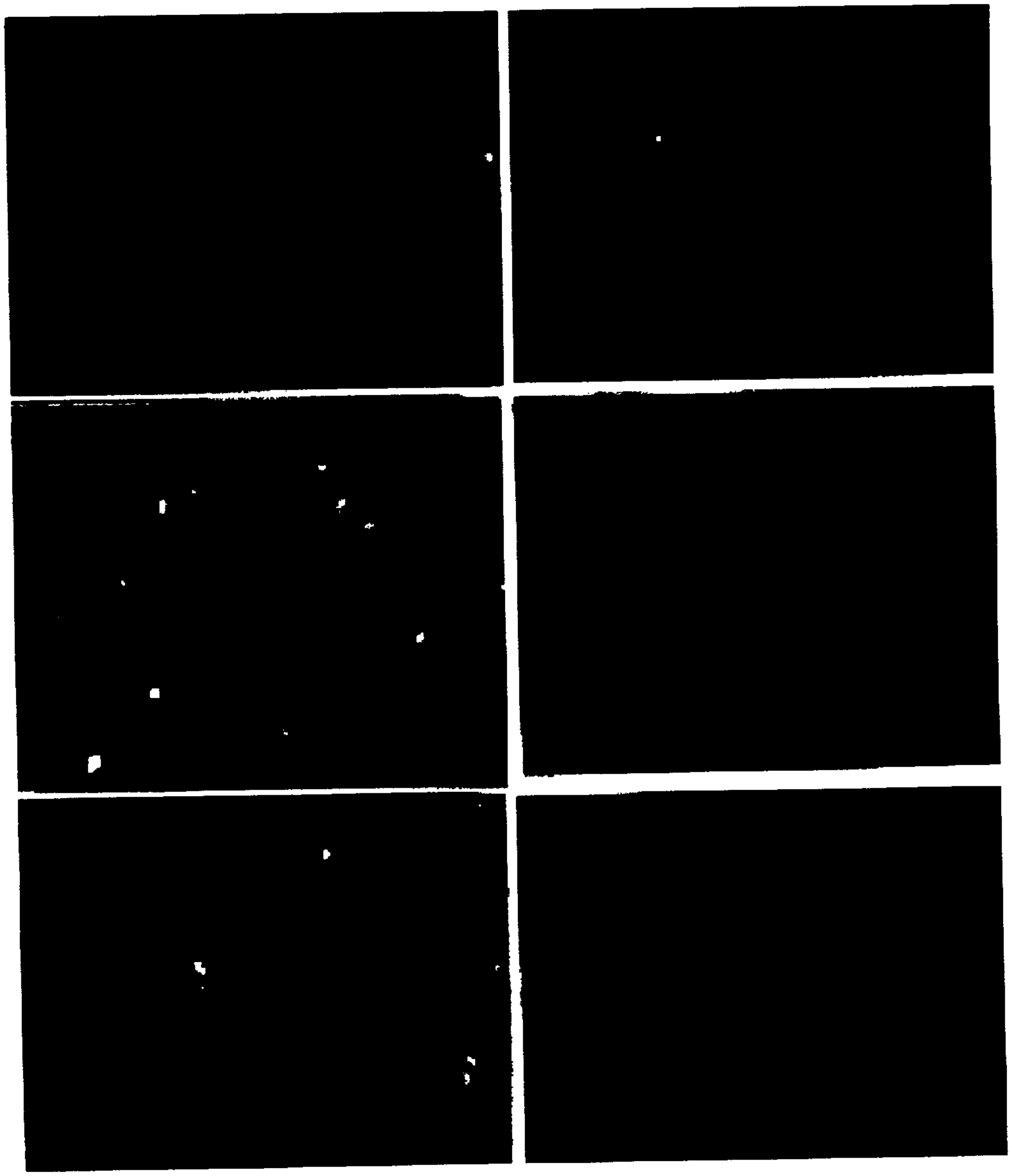


Figure 1. Flourescence micrographs of BAEC after 24, 48 and 72 hr under control (A, C & E) and 12G hypergravity (B, D & F) conditions. Cultures are stained for f-actin using FITC-labeled phalloidin.

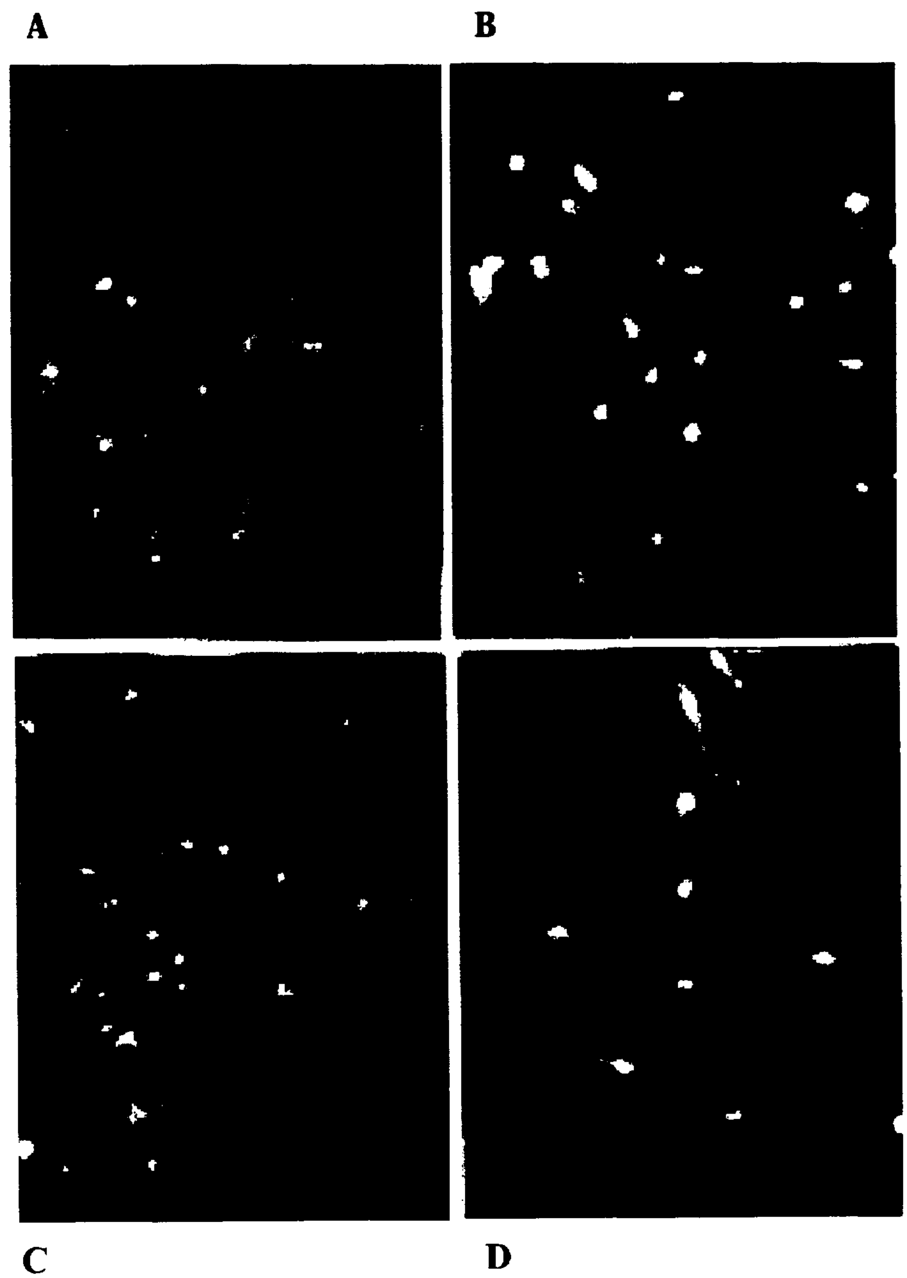


Figure 2. Flourescence micrographs of BAEC treated with PKI KT5926 (for calmodulin(I I)-dependent protein kinase) after 24 and 48 hr under control (A & C) and 6G hypergravity (B & D) conditions. Cultures are stained for f-actin using FITC-labeled phalloidin.

STUDIES ON PTCDA/NTCDA MULTIPLE QUANTUM WELLS

C. W. Lowe, S. C. Mathur, Renuka Verma*, and Kang I. Seo
Research Center for Optical Physics, Department of Physics,
Hampton University, Hampton, VA 23668

INTRODUCTION

In order to realize effective optical information processing, we should find out nonlinear optical materials with large third-order optical polarizability $\chi^{(3)}(-\omega; \omega, -o), \omega)$ and short switching time τ at the same time. When one defines the figure-of-merit for nonlinear optical response F by $F = |\chi^{(3)}| / \alpha(\omega)\tau$, F was found[1] to be almost constant, independent of materials and pump frequency. Here $\alpha(\omega)$ is the linear absorption coefficient at relevant frequency ω. As a result, the material with fast response time shows a small $\chi^{(3)}$. This is the case with organic solids, $\chi^{(3)}$ up to 3×10^{-6} esu has been reported for thin films of polyacetylene[2]. This value lies far beneath these displayed by semiconductor-based multiple quantum wells; however, within the transparency domain, organic materials exhibit a much faster response. If we remember that $\chi^{(3)}$ for AlGaAs/GaAs multiple quantum well (MQW) is much larger that GaAs, we may hope organic multiple quantum wells (OMQW) to show $\chi^{(3)}$ values much larger than organic semiconductors. Thus OMQW may he the structures with improved figure-of-merit F. For OMQW lattice mismatch is not a problem due to weak vander-waal binding forces between molecules in an organic solid. This gives a wide choice of materials to choose components of OMQW. So et al.[3] and Ammermann et al.[4] have reported OMQW structures. Ammermann et al. report the fabrication organic light emitting diode (OLED) using OMQW based on tris-(8-hydroxy-chinoline) aluminum/2-(4-biphenyl)-5-(4-tert-butylphenyl)-1,3-4-oxadiazole. So et at. fabricated OMQW using 3,4,9,10 perylene tetra-carboxylic di-anhydride (PTCDA) and 3,4,7, 8 naphthalene tetra-carboxylic di-anhydride (NTCDA).

A Quantum well consists of an ultrathin layer of a lower bandgap semiconductor sandwiched between two lattice matched, wide band-gap semiconductors. For OMQW one should select two organic semiconductors with sufficient difference between bandgaps with lower bandgap region being inside the wider band-gap region. Therefore, one may like to do energy band structure calculations for organic crystals. However, a simpler approach to selection of materials for OMQW is provided by highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energies. HOMO and LUMO energies lie at the center of energy bands. We have used this approach to look into the OMQW based on PTCDA and NTCDA.

CALCULATIONS

We have carried out simple LCAO-MO calculations for PTCDA and NTCDA molecules. For the calculations the values of coulomb integral (α) and resonance integral (β) have been taken to be -5.96 eV and -3.08 eV respectively. These values are based on empirical correlation of ionization potential and spectra of a number of organic molecules[5,6]. The hetero-atomic nature of the PTCDA and NTCDA molecules has been taken following Streitweiser[6]. Calculated HOMO energies and some of the higher-state energies have been collected in table 1, Only those higher-state energies have been included in table 1 for which spectral transitions are above 200 nm (limit of available experimental data).

RESULTS AND DISCUSSIONS

As can be seen from table 1, the HOMO-LUMO energy separations for PTCDA and NTCDS are 1.49 eV and 0.72 eV respective y. Because HOMO and LUMO energies lies at the center of energy bands, the energy separation between HOMO and LUMO may be approximated to energy band gaps. Based on their spectral studies So et al.[3] have reported a band-gap of 2.2 eV for PTCDA and 3.1 eV for NTCDA. Theoretical and experimental values of band gap for PTCDA are in reasonably good agreement if one remembers that we have done only semi-empirical calculations using parameter values based on hydro-carbon molecules. However, the calculated band gap of NTCDA is much smaller than experimental value of the band gap. Further this band gap is smaller than the calculated band gap for PTCDA while the experimental values show that NTCDA bandgap is larger than the PTCDA bandgap. A look at the theoretical spectral transitions suggest that the transition between HOMO and LUMO lies at 1719 nm. This value is outside the range of spectra reported by So et al.[3] According to our calculated values next higher spectral transition is at 403 nm. This value corresponds

to 3·08 eV and is equal to the band-gap reported by So et al. Since So et al.[3] did not extend the range of their spectra, they took the -400 nm peak to represent the energy band-gap of NTCDA.

The spectra reported by So et al. is not adequate to examine properly. So we recorded the spectra of PTCDA and NTCDA in thin film form. These spectra are shown in figure 2. The values of peak positions are collected in table 1. Our spectra is quite similar to the ones reported by So et al. As can be seen from table 1, the agreement between theoretical and experimental values of peak positions of NTCDA spectra is extreemly good. Our calculated peak positions are at 403 nm, 268 nm, and 229 nm. Corresponding experimental values are 397 nm, 249 nm, and 225 nm. However in case of PTCDA the agreement is not so good. For PTCDA we find the lowest energy peak is at 570 nm according to our spectra while the calculated value is at 836 nm. But the theoretical values of 258 nm and 229 nm are in an excellent agreement with experimental values of 253 nm and 226 nm. This gives us confidence in our calculations.

To visualize the organic multiple quantum wells (OMQW) based on PTCDA/NTCDA we have used the values of table 1 and drawn an energy level diagram shown in figure 1. In the figure the LUMO energies have been taken as the bottom of conduction bands and HOMO energies as top of the valance bands. This energy diagram of OMQW based on PTCDA/NTCDA is quite different from the well known AlGa As/ Ga As energy diagram. In the PTCDA/NTCDA OMQW the wells are quite deep but there is no step potential in the valance band. However, it is clear from these calculations that the energy gap of PTCDA is larger than that of NTCDA and that the NTCDA band-gap region lies inside the band-gap region of PTCDA. Therefore, these two materials are suitable for OMQW. Similar studies on other organic materials may bring out the pairs of organic materials suitable for OMQW. Calculation of energy states of PTCDA and NTCDA thin films are in progress and will be reported soon.

References:

1. D. H. Auston, T.K. Gustafson, A. E. Kaplan, P. L. Kelley and Y. R. Shen: Appl. Opt. 26, 211 (1987)
2. F. Krausz, E. Winter and G. Leising: Phys. Rev. B39, 3701 (1989)
3. F. F. So, S. R. Forrest, Y. Q. Shi and W. H. Steier: Appl. Phys. Lett. 56, 674 (1990)
4. D. Ammermann, A. Böhler, C. Rompf and W. Kowalsky: Presented at ACS meeting on OrganicThin Films for Photonic Applications held at Portland, OR (1995)
5. S. C. Mathur and D. C. Singh: Ind. J. Pure Appl. Phys. 8, 788 (1970)
6. A. Stretwieser Jr.: Molecular Orbital Theory for Organic Chemists, *John Wiley & Sons, Inc.* (1961)

TABLE I

Energy eigen Value (in eV)	Energy separation from HMO (in eV)	Spectral Peak position λ (in nm) Calculated	experimental
		PTCDA	
-9.04 (HOMO)			
-7.55 (LUMO)	1.49	836	553
			503
			480
-5.96	3.08	403	391
-4.37	4.67	266	
-4.22	4.82	258	253
-3.62	5.42	229	226
		(NTCDA)	
-9.04 (HOMO)			
-8.32 (LUMO)	0.72	1719	
-5.96	3.08	403	393
-4.41	4.63	268	249
-3.60	5.44	228	225

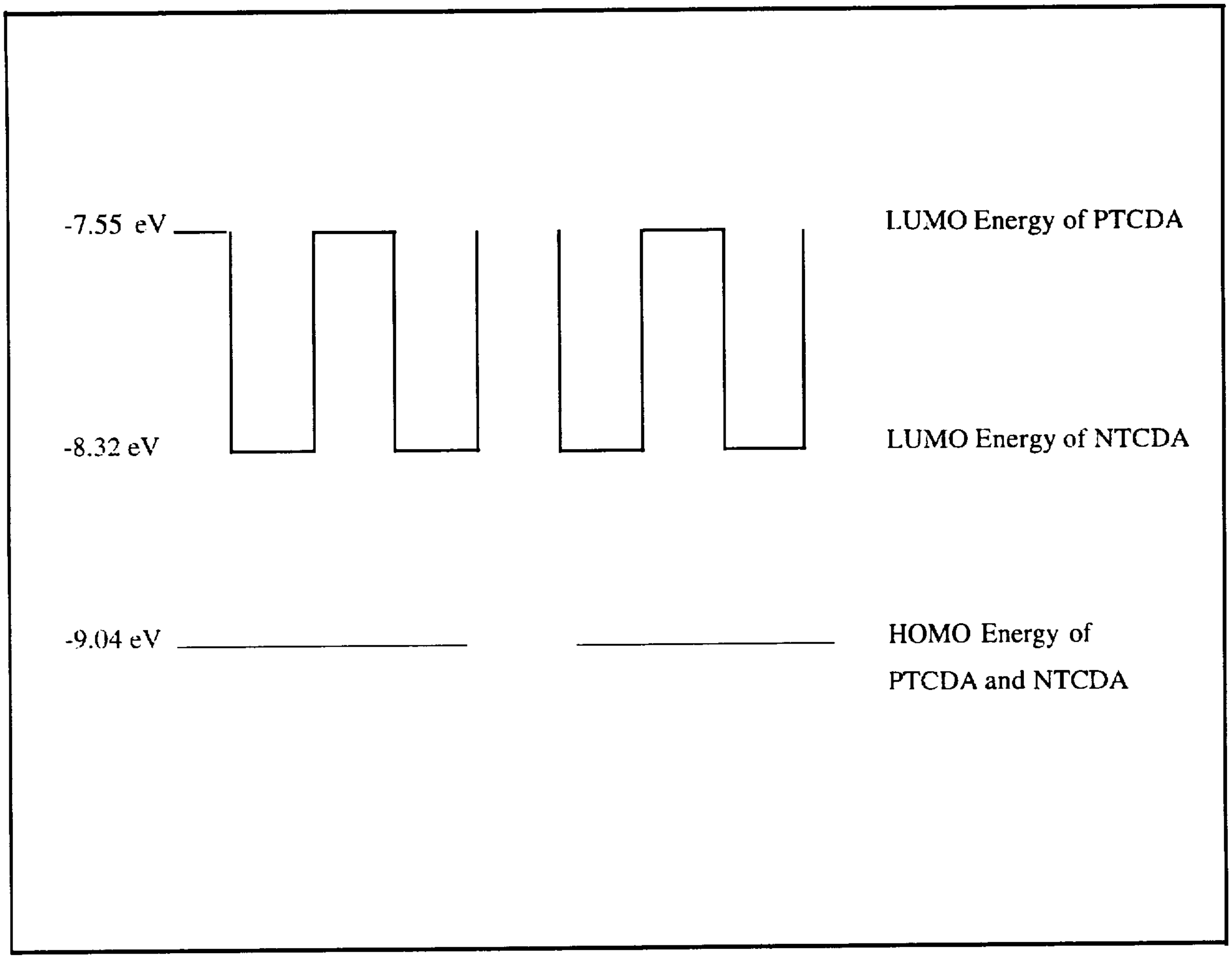


Figure I: Energy level diagram for PTCDA/NTCDA Multiple Quantum Wells

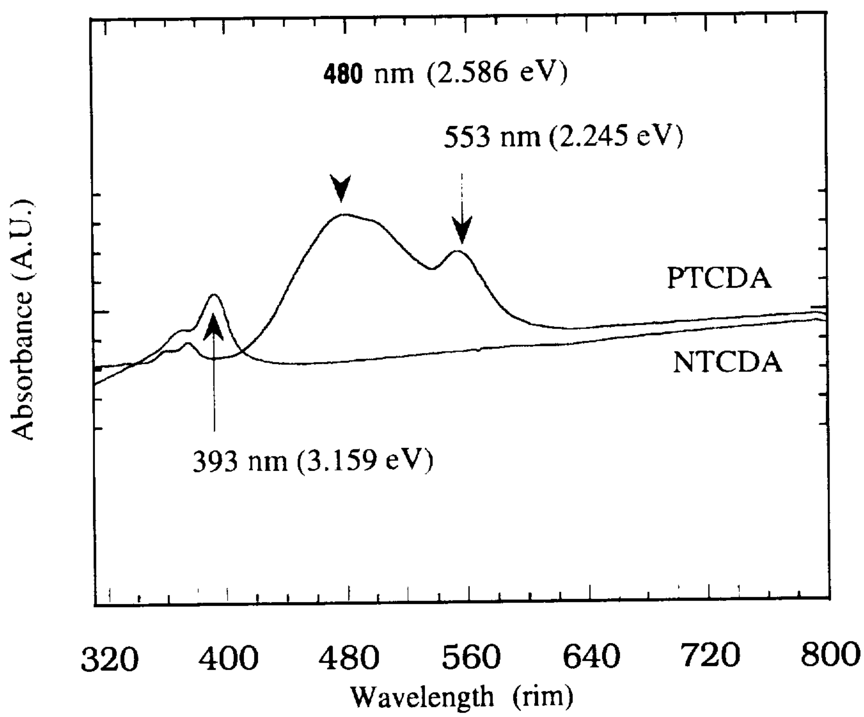


Figure 2. Spectra of PTCDAand NTCDA thin films.

Page intentionally left blank

URC97151

Potentials for Soil Enzyme Activities as Indicators of Ecological Management

Z. N. Senwo*, A. Manu, and T. L. Coleman
Center for Hydrology, Soil Climatology, and Remote Sensing
P.O. Box 1208
Alabama A&M University
Normal, AL 35762

Abstract

Activity measurements of selected soil enzymes (cellulase, glucosidase, amidohydrolase, phosphatase, arylsulfatase) involved in carbon, nitrogen, phosphorus, and sulfur cycling in the biosphere, hold potential as early and sensitive indicators of soil ecological stress and restoration, These measurements are advantageous, because the procedures are simple, rapid, and reproducible over time. Enzyme activities are also sensitive to short-term changes in soil and kind-use management. Enzyme activities have also been observed to be closely related to soil organic matter proposed as an index of soil quality.

Introduction

Public concerns about soil, water, and environmental degradations have increased significantly and agricultural practices are criticized as a major contributor. The activity measurements of some selected soil enzymes (cellulase, α- or β- glucosidase, amidohydrolase, acidic or alkaline phosphatase, and arylsulfatase) involved in carbon, nitrogen, phosphorus, and sulfur cycling in the biosphere, are being considered as potential indicators of soil management practices, soil quality/health, ecological stress and restoration. Soil enzyme activities have recently been considered among the most efficient and cost-effective tools for analyzing changes in various land management practices such as residue management, soil compaction, tillage, and crop rotation (Dick, 1994, Deng, 1994, Senwo, 1995),

Soil enzyme activities can be used: (1) to test and/or generate hypotheses to improve understanding of soil biological, chemical, and physical processes; (ii) to provide guidelines in identifying gaps in knowledge and stimulate new research initiatives; (iii) to integrate basic knowledge of various biochemical and physical attributes of the biosphere; (iv) to make long-term predictions of the impacts of agricultural practices on soil, water, and environmental quality; and (v) to select the best alternative practices fitting the desired soil, water, and environmental quality goals.

Enzyme activities in soils have been shown to be closely related to other proposed indexes (organic matter, pi-Q of soil quality (Table 1).

Table 1. Correlation coefficients for linear regressions of enzymatic activities and organic C or pH of soil under tillage and residue management.

	Correlation Coefficient	
Enzyme activity	Organic C	pH
	r[†]	
Aspartase	0.84***	0.41**
Amidase	0.90***	0.24
L-asparaginase	0. SO***	0.74***
L-glutaminase	0.70***	0.77***
Urease	0.80***	0.72***

†**, ***, Significant at P <0.01, and 0.001 respectively (From Senwo, 1995).

Soil organic matter decomposes very slowly and many years may be required to measure changes from decomposition activities. The accumulation of organic and inorganic nutrients in

soils stimulates microbial growth and activity, and therefore, enzyme synthesis. High organic matter levels from residue applications may provide more favorable environment for the accumulation of enzymes in the soil matrix (Burns, 1982). Enzymes in soils may be polymerized, entrapped, and/or adsorbed giving rise to a stable active enzyme-soil colloid associations (Burns, 1982) and their activities in soils are also closely related (Table 2).

Table 2. Correlation coefficients for linear regressions of between enzyme activities.

	Correlation Coefficient			
	Aspartase	Cellulase	α-Gluco	β- Gluco
	$r^{\dagger}$			
Arylsulfatase	na	0.33*	0.74***	0.46**
Amidase	0.44**	0.66***	0.67***	0.61***
L-asparaginase	0.94***	0,43**	0.76***	0.53***
L-glutaminase	0.88***	0.40**	0.66***	0.48**
Urease	0. SO***	0.40**	0.87***	0.50**

†*,**,***, Significant at $P < 0.05$, 0.01, and 0.001 respectively. (From Deng, 1994; Senwo, 1995). na = not available, Glu = glucosidase.

Various soil management practices have profound effects on enzyme activities (Dick, 1994; Deng, 1994; Senwo, 1995). Gupta and Germida (1988) observed that cultivation depressed phosphatase and arylsulfatase activities by 49 and 65%, respectively. Senwo (1 995) observed aspartase activity in soils were affected by tillage and management practices (Fig. 1).

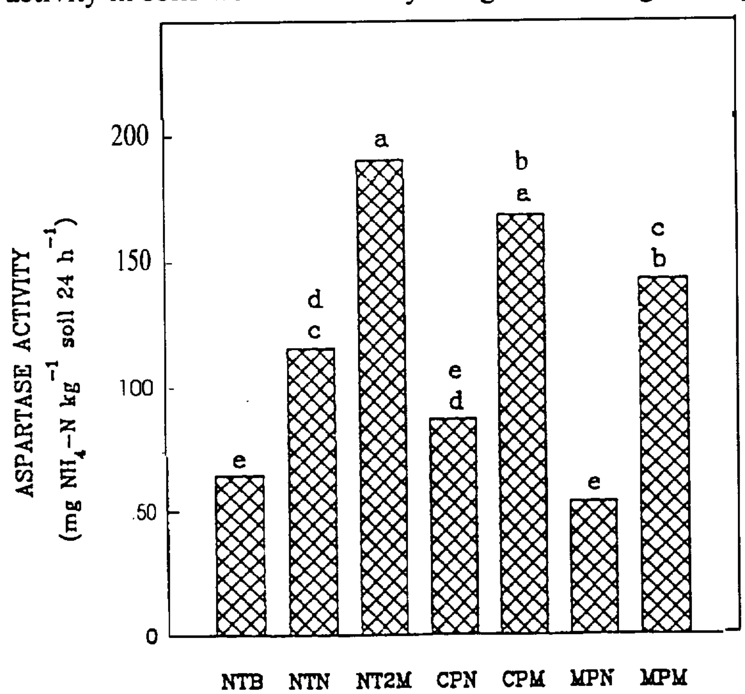


Fig. 1. Bars with the same letter are not significantly different ($P < 0.05$). NTB = not tilled but bared, NTN = not

tilled but not bared, NT2M = not till but doubled mulch, CPN = chisel plowed and not mulched, CPM= chisel plowed and mulched, MPN = moldboard plowed and not mulched, MPM = mold board plowed and mulched (adapted from Senwo, 1995).

Deng (1994) also reported that the activities of 14 enzymes involved in C, N, P, and S cycling in soils were greater in four replicated plots that were not tilled but doubled mulched, than in those treated with other tillage systems and residue placement. The activities decreased significantly with increasing soil depth, accompanied by a decrease in organic C content and pH. Eivazi and Bayan (1994) observed that the activities of et- and β-glucosidase, and acid phosphatase were significantly reduced by burning treatments.

Activity Measurement

There has been little or no success in extracting enzymes from soils (Tabatabai, 1982), however, several procedures exist for measuring enzyme activities in soils (Tabatabai, 1994; Alef and Nannipieri, 1995; Senwo and Tabatabai, 1996). The measurement of most soil enzyme activities involve the quantitative measurement of the appearance or disappearance of a product when soil has been treated with a microbial inhibitor (usually toluene) and incubated with buffered solution at a fixed temperature and time. The pertinent parameters measured for most enzyme assays include: the optimal pH (Fig. 2), "

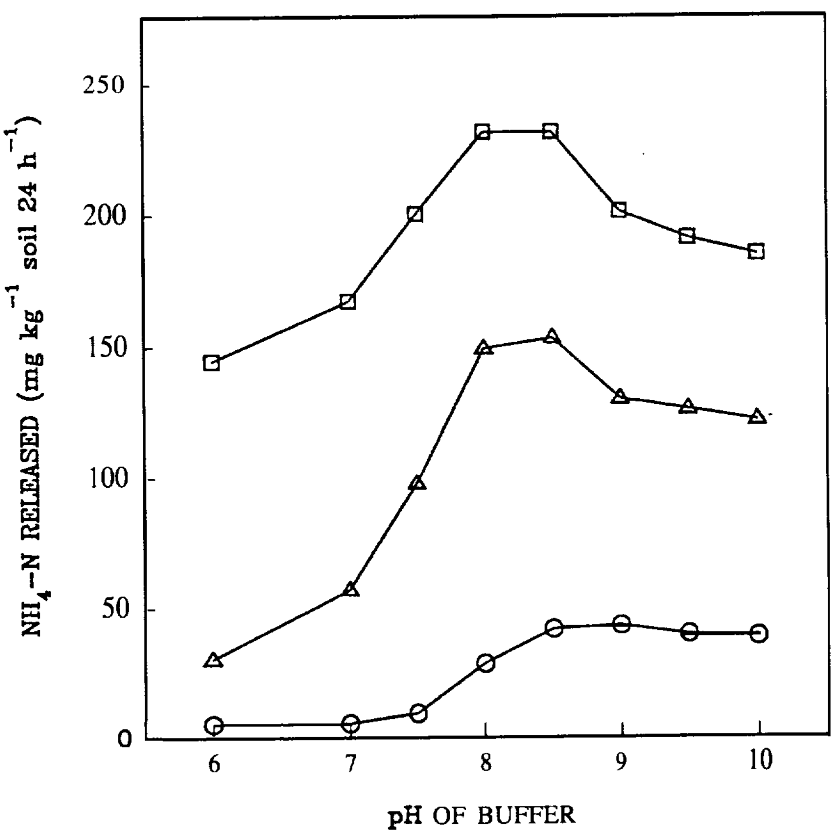


Fig. 2. Effect of pH of buffer on aspartase activity in soils. (From Senwo and Tabatabai, 1996).

substrate concentration at which the reaction essentially follows a zero-order kinetics, amount of soil needed to obtain maximum activity without limiting the substrate concentration, temperature

(Fig. 3) and time of incubation to obtain maximum activity.

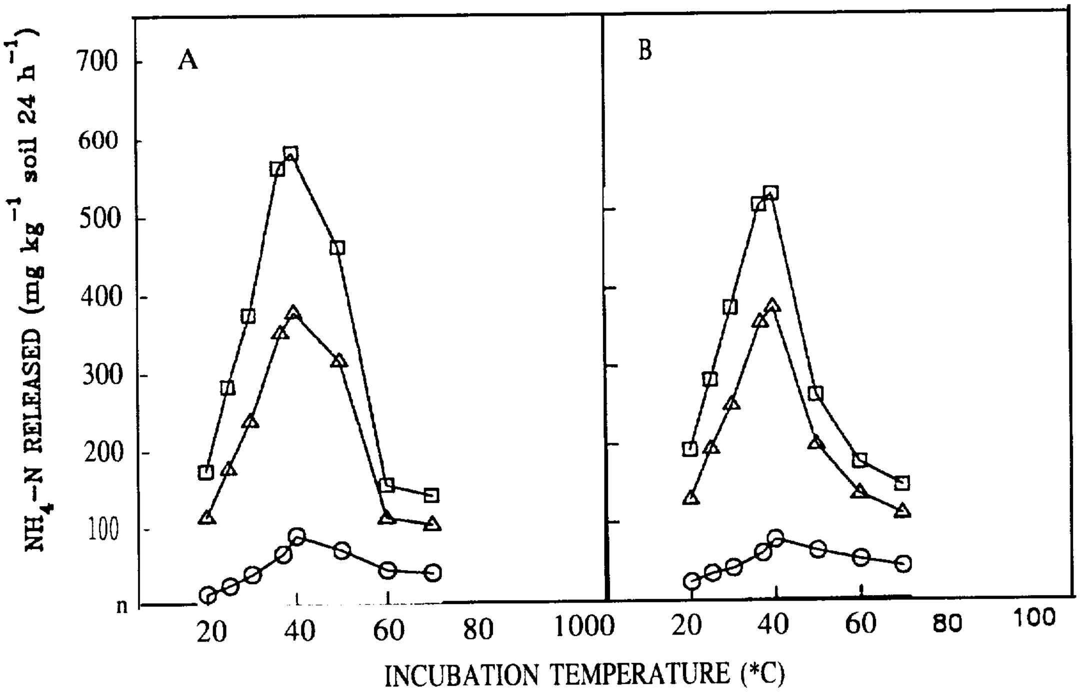


Fig. 3. Effect of incubation temperature on aspartase activity in soils. A, Field-moist soils; B, air-dried soils (From Senwo and Tabatabai, 1996).

Summary and Conclusion

The potential exist for use of soil enzyme assays in identifying positive or negative effects of land management practices within periods, long before there are measurable changes in soil organic matter. Most soil enzyme assays are simple, rapid, and reproducible.

Acknowledgments

Acknowledgment is extended to the Center for Hydrology, Soil Climatology, and Remote Sensing (HSCaRs) support staff. Contribution from HSCaRS, the Department of Plant, Soil and Animal Science, and the Alabama Agricultural Experiment Station, Alabama A&M University, Normal, AL 35762. Journal No. 347.

This paper was supported by Grant No. NCCW-0084 from the National Aeronautics and Space Administration (NASA), Washington, DC. Any use of trade, product or firm names is for descriptive purposes only and does not imply endorsement by the U. S. Government.

References

Alef, h'., and P. Nannipieri. 1995. Methods unapplied soil microbiology and biochemistry. Academic Press, San Diego, CA.

Burns, R.G. 1982. Enzyme activity in soil: Location and a possible role in microbial ecology. Soil Biol. Biochem. 14:423-427.

Deng, S.P. 1994. Cellulose activity of soils and the effect of tillage management on enzyme activities in soils. Ph.D. dissertation, Iowa State University,

Dick, R.P. 1994. Soil enzyme activities as indicators of soil quality. p. 107-124. *In* Doran et al. (cd.) Defining soil quality for sustainable environment, SSSA Spec. Publ. No. 35. SSSA and ASA, Madison, WI.

Eivazi, F., and M.R. Bayan. 1994, Soil enzymes activities and microbial biomass in an oak-Hickory forest following long-term prescribed burning. Agron. Abst. Madison, WI.

Gupta, V. V. S. R., and J.J. Germida. 1988. Distribution of microbial biomass and its activity in different soil aggregate size classes as affected by cultivation. Soil Biochem. 20:777-786.

Senwo, Z.N. 1995. Amino acid composition of soil organic matter and nitrogen transformations in soils under different management systems. Ph.D. dissertation, Iowa State University.

Senwo, Z.N., and M.A. Tabatabai. 1996. Aspartase activity of soils. Soil Sci. Sot. Am. J. 60:1416-1422.

Tabatabai, M.A. 1982. Soil enzymes. p. 903-945. *In* A.L. Page et al. (cd.). Methods of soil analysis. Part 2. 2nd ed. Agron. Monog. 9. ASA and SSSA, Madison, WI.

Tabatabai, M.A. 1994. Soil enzymes. *In* R.W. Weaver et al. (cd.) Methods of soil analysis, Part 2, microbiological and biochemical properties. SSSA Book Series, 5:775-833.

Page intentionally left blank

/ / / / / / / / /

URC97152

Hypergravity Alters the Susceptibility of Cells to Anoxia-Reoxygenation Injury

Henry McCloud*, Yulondo Pink[§], Sandra A. Harris-Hooker[‡], Caroline D. Melhado[§] and Gary L. Sanford[§]

*Department of Biology, Morehouse College
and
Departments of Biochemistry[§] and Medicine[‡]
Morehouse School of Medicine
Atlanta, Georgia 30310.

INTRODUCTION: Gravity is a physical force, much like shear stress or mechanical stretch, and should affect organ and cellular function. Researchers have shown that gravity plays a role in ventilation and blood flow distribution, gas exchange, alveolar size and mechanical stresses within the lung (1-3). Short exposure to microgravity produced marked alterations in lung blood flow and ventilation distribution while hypergravity exaggerated the regional differences in lung structure and function resulting in reduced ventilation at the base and no ventilation of the upper half of the lung (4). Microgravity also decreased metabolic activity in cardiac cells, WI-38 embryonic lung cells, and human lymphocytes (5). Rats, in the tail-suspended head-down tilt model, experienced transient loss of lung water (6), contrary to an expected increase due to pooling of blood in the pulmonary vasculature. Hypergravity has also been found to increase the proliferation of several different cell lines (e.g., chick embryo fibroblasts) while decreasing cell motility (7) and slowing liver regeneration following partial hepatectomy (8). These studies show that changes in the gravity environment will affect several aspects of organ and cellular function and produce major change in blood flow and tissue/organ perfusion. However, these past studies have not addressed whether ischemia-reperfusion injury will be exacerbated or ameliorated by changes in the gravity environment, e.g., space flight. Currently, nothing is known about how gravity will affect the susceptibility of different lung and vascular cells to this type of injury. We conducted studies that addressed the following question: Does the susceptibility of lung fibroblasts, vascular smooth muscle and endothelial cells to anoxia/reoxygenation injury change following exposure to hypergravity conditions?

EXPERIMENTAL METHODS: Bovine aorta endothelial (BAEC) and primate smooth muscle(SMC) were obtained from the NIGMS/Coriell Cell Repository. Rat lung fibroblasts (RFL) were isolated from adult Sprague Dawley rats as previously reported (9). All cell lines were maintained in DMEM containing 10% fetal bovine serum. Confluent cultures of each cell line were subjected to centrifugation at 6G for 24-48 hrs. Control cultures were not centrifuged or rotated. Cells were then placed under anoxia (5% CO_2 /balance N_2) for 2 hr with or without a 1 hr period of reoxygenation. The change in viable cell numbers was assessed by: measuring viable cells by hemacytometer counting of trypan blue stained cells, or using MTT assay for viable cells (microtiter plate assay). The effect of hypergravity on the expression of heat shock protein (HSP60) by RFL was evaluated by immunocytochemical staining using a FITC-labeled monoclinal antibodies (Stress Gene). Subconfluent cultures were fixed after 12 and 24 hr under hypergravity using 10%formalin, 0.1 % triton X-100 in PBS. Cultures were incubated with the primary antibody for 1 hr, washed 3X with PBS, incubated with biotin labeled anti IgG secondary antibody for 1 hr and subsequently stained with fluorescein conjugated streptavidin. Cells were viewed by fluorescence microscopy and photographed. Controls, stained with non-immune mouse IgG, showed no fluorescence under these conditions.

RESULTS AND DISCUSSION: Figure 1 shows the change in viable cell count, expressed as percent of controls, for SMC and BAEC as a function of time under 6G. The solid line at 100°/0 represents the level for cultures that were not treated with anoxia or anoxia-reoxygenation. SMC were found to have a 60°/0 decrease in viable cells after 48 hr of hypergravity and 2-hr of anoxia. In the first 24 hr of hypergravity, the acute response of SMC seems to be an exacerbated injury under anoxia and reoxygenation. This acute phase is followed in the next 24 hr by a reversal of the exacerbated injury as cells adapt to hypergravity.

The results found for BAEC show a different pattern. Control (non-centrifuged) cells had decreased viability under anoxia with only a slight further decrease during the reoxygenation period. With BAEC, the acute response (first 24 hr) to hypergravity is an increase in viable cells under anoxia-reoxygenation. As seen with SMC, there is a reversal of this increase in the following 24 hr period

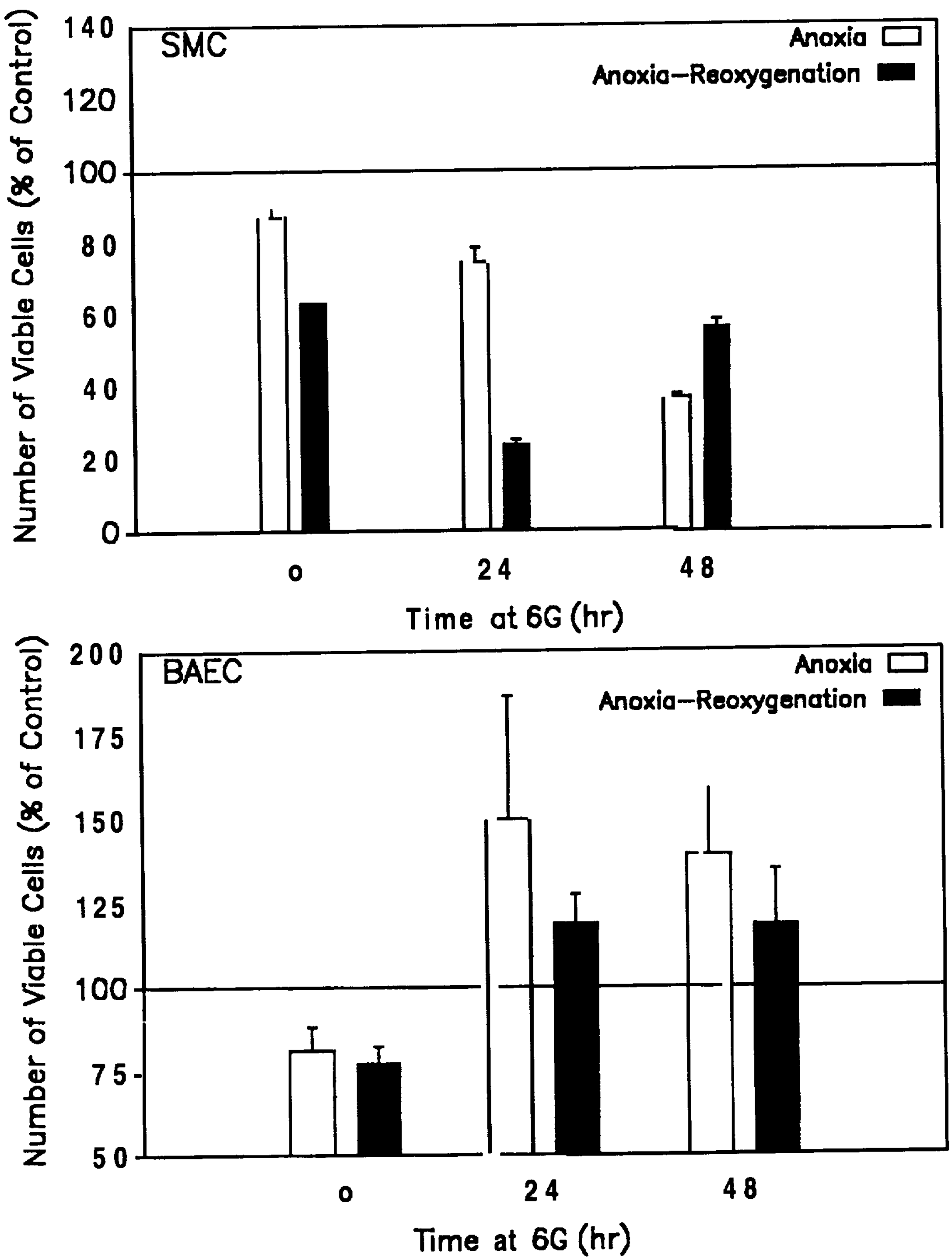


FIGURE 1. The response of SMC and BAEC to anoxia or anoxia-reperfusion injury following hypergravity treatment for 24-48 hr. Data is plotted as mean of the % of matched untreated (controls) normogravity or hypergravity cultures.

as the cells adapt to hypergravity. Although it is expected that more cell damage and loss of viability would result from the combined effect of anoxia and reoxygenation, we found that 2 hr of anoxia results in similar damage. Hypergravit y worsened the damage for SMC but decreased the damage for BAEC, suggesting that protective mechanisms maybe differentially activated in vascular cells.

Figure 2 shows that results similar to those seen with BAEC were also found for RFL whether these cells were maintained in DMEM or PBS. When maintained in complete media (DMEM), these cells show an increase in viable cells with time in hypergravity when treated with anoxia alone. However, RFL maintained in PBS show a steady decrease with time in hypergravity. RFL behave similar to BAEC when placed under anoxia-reoxygenation. The acute response of RFL to hypergravity is a transient decrease in susceptibility to damage, i.e., increased viable cell counts, under anoxia-reoxygenation. Again this is followed by a reversal in the following 24 hr period as the cells adapt to hypergravity. This pattern of response was also found for RFL maintained in PBS.

When BAEC and RFL have apparently adapted to hypergravity, both cell lines are not readily injured by anoxia-reoxygenation. In fact, RFL maintained in PBS, which has no energy source (e.g., glucose) was not found to have the maximum injury under hypergravity as expected. Under these conditions RFL responded similarly to cells in complete media. These results suggest that hypergravity may increase protective mechanisms in these cells that ameliorate possible damage from anoxia-reoxygenation. One possible protective mechanism that maybe induced by hypergravity is increased expression of stress proteins like heat shock proteins. Heat shock proteins have been shown to protect cells from damage and death under a number of different stresses, e.g., increased temperature or sheer stress (10). We tested this possibility with RFL, as shown in figure 3. Within 12 hr under hypergravity, RFL's had clearly increased immunofluorescence for heat shock protein, HSP60 (figure 3A) compared to control cells (figure 3 C). By 24 hr, hypergravity resulted in an even more intense immunofluorescence, suggesting that HSP60 expression is stimulated for at least the initial 24 hr period. Preliminary data from similar studies with BAEC (data not shown) indicate that these cells also may increase their expression of heat shock proteins within the first 24 hr period of hypergravity. Further studies are planned to examine the possible mediation of hypergravity effects on cells by the heat shock protein family.

ACKNOWLEDGEMENTS: These studies were supported by grants from NASA NAG9-644 and NCCW-0085.

REFERENCES

1. Prisk GK, et al., J. Appl. Physiol. 78:597-607, 1995.
2. Elliott AR, et al., J. AppL. Physiol. 77:2005-2014, 1994.
3. Prisk G K, et al., J. Appl. Physiol. 75:15-26, 1993.
4. West JB, et al., The Physiologists 25, S21-S24, 1982.
5. Cogoli A, The Physiologists 28, S47-S50, 1985.
6. Hargens AR, et al., The Physiologists 28, S155-S156, 1985.
7. Tschopp A and Cogoli A, Experiential 39, 1323-1329, 1983.
8. Kropacova K, et al., The Physiologists 31, S75-S76, 1988.
9. Mary B. Jones et al., Am. Rev. Resp. Dis., 135:997-1001, 1987,
10. Healy AM, et al., Annals NY Acad. Sci. 663:319-330, 1992.
11. Bernelli-Zazzera, et al., Annals NY Acad. Sci. 663:120-124, 1992.

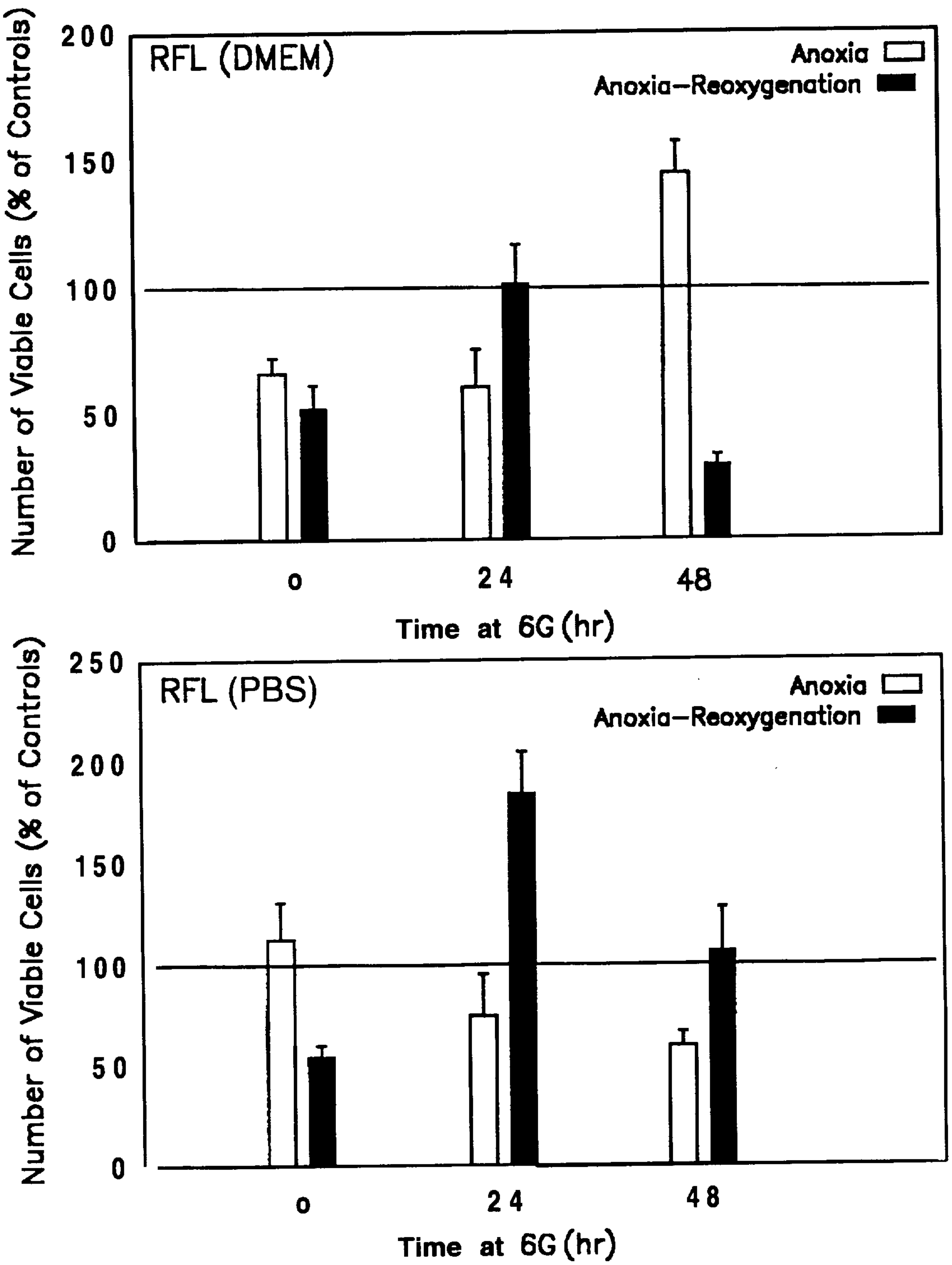


FIGURE 2. The response of RFL, maintained in complete media (DMEM) or in PBS, to anoxia or anoxia-reperfusion injury following hypergravity treatment for 24-48 hr. Data is plotted as mean of the 1%0 of matched untreated (controls) normogravity or hypergravity cultures.

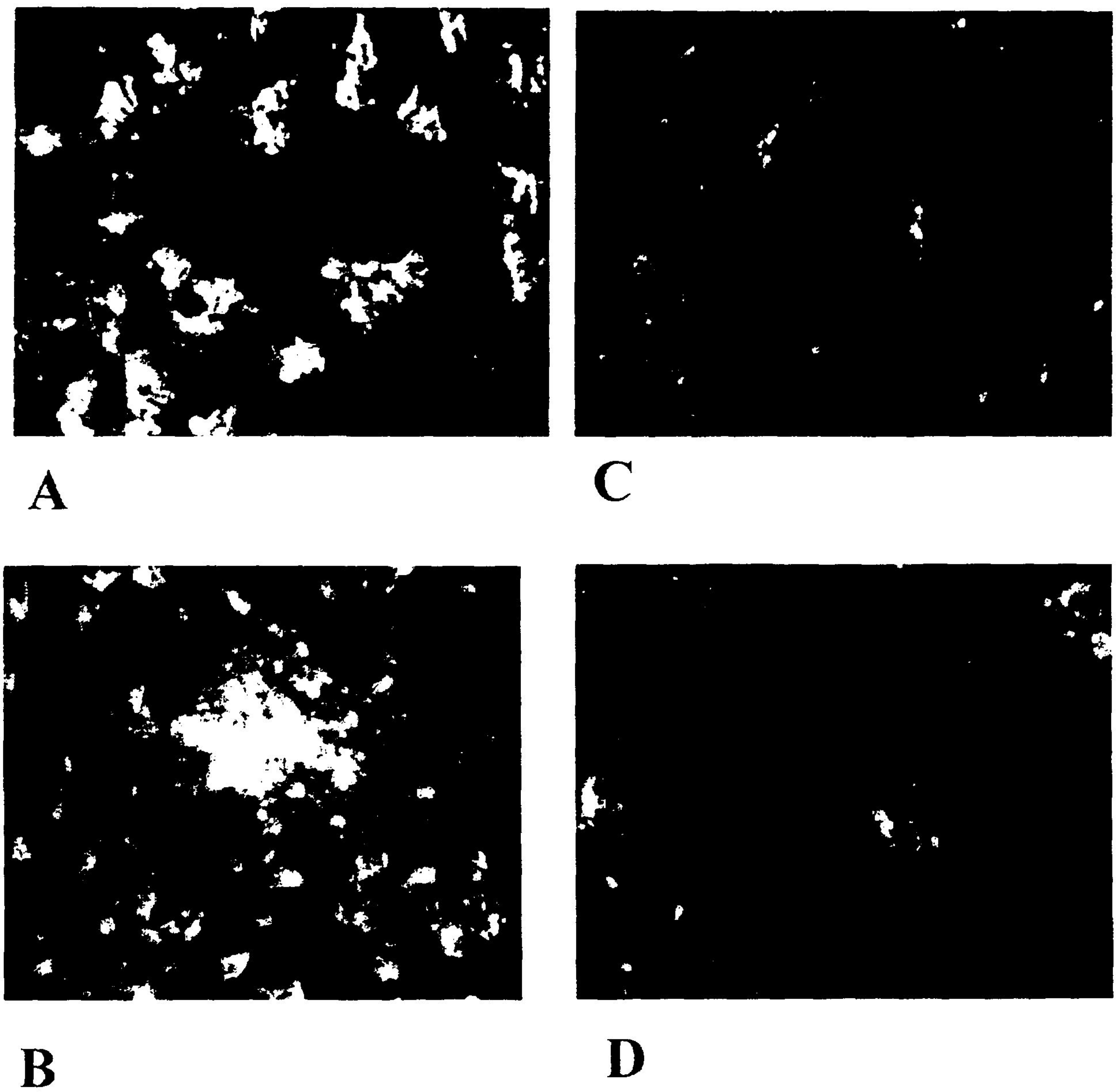


Figure 3. Expression of heat shock protein (HSP60) by RFL subjected to 12 and 24 hr of h ypergravity (A & B) or normogravity (C & D).

URC97153

On Establishing a Connection Between Fuzzy Relations and Linear Block Codes

Roger Salters
Associate Professor
Department of Engineering
University of Denver
Devner, CO

Marisol Gamboa
Student, Computer Science
University of New Mexico
Albuquerque, NM

ABSTRACT A binary linear block code is shown to be isomorphic to a fuzzy relation matrix under a specified λ-cut. This paper provides the details which establish this connection between fuzzy relations and linear block codes. The fuzzy relation matrices are developed from the time-series data using outer vector products. Then directed graphs and trellises are defined, and, finally, the edge-minimal trellises produce the linear block codewords.

L Introduction

We consider a new problem in defining fuzzy relations on crisp or fuzzy time-series data where the results are a continuous message of the codewords equivalent to the data. These data points are multidimensional and the maps of these points evolve sequentially from some starting point, say LO, to a final point, say L_f. The process of evolving the characteristics or attributes of the data at L. to any intermediate point L_i in the associated directed graph shows that each path through the graph can be continuous and the influence at coordinate values at LO reaches each coordinate value at L_t As will be shown later, this sequential coupling of characteristics or attributes through the evolving time-series can be represented by a continuous-time variable memory Markov process. To the author's knowledge, the only fuzzy relations that has a similar property of memory is the time-varying decomposable relation (d-relation), where, given the characteristic properties of the universal set, one may decompose or partition the universal set into subsets conditioned on that particular time-varying property [1]. The conditioning can be time-varying. These conditional subsets support pseudomeasures that provide topological structure on the universal set. Such is also the case for our relations. According to the Literature, no other fuzzy relations are endowed with pseudomeasures as part of their definitions [1], [2], [3], [4], [5], [6]. Thus the topological structure that is inherent in the data appears not to be included as an essential basis for defining the relations. (In Section III, we develop such a relation.) However, in [2, Chapter 14], [7] fundamental tools are presented on which families of unique results may be formed, but our method presents as its product the unique family results. In the approach presented in this paper, the optimization of the relation to the particular multidimensional data is inherent in the results. The related or associated graph of the relation, the edge-minimal trellis, and the resulting binary linear block codewords reflect the uniqueness of the optimization process inherent in the final code set

The rest of the paper is organized as follows, In Section II, we shall present preliminaries of each the fundamental used in the paper. These discussions are brief but references are made to additional sources in the literature. We also present a description of the multidimensional time-series date which motivated the paper. Section III contains our main results. Here we bring together in a coherent form the preliminaries from Section II. We also present some observations and comments on the results in this section. The paper is concluded in Section IV.

II. Preliminaries

In this section we present background materials on essential concepts that are used in the sequel. Several of these concepts are given as definitions and others are given in terms of notation used in Section III. first, we will discuss directed graphs and relations, second, directed graphs and trellises, and finally trellises and linear block codes.

A. Directed Graphs and Relations

A directed (weighted) graph [8] is a pair $(V,E) = G$, such that the set of vertices $V=V(G)$ and the set of edges $E=V(G)$ show that these vertices and edges are in graph G.

If vertices u and v are in V(G) and edge $e=e_{uv}$ is in E(G), then we say that edge e of the graph connects vertex u to vertex v. Likewise, we can define a relation associated with graph G as

(1) $\gamma_{uv} \geq O$ for every $u,v \in V(G)$ ①

(2) $\gamma_{uv} \geq O$ iff $e_{uv} \in E(G)$ and $e_{uv} = e_{vu}$ for every $u, v \in V(G)$

(3) The cardinality of V(G) is $n=n(G)=|V(G)|$ for a finite graph.

NOTE: For a fuzzy graph the cardinality is infinite [2].

For the binary graph, that is, a normalized graph, for $e_{uv} \in E(G)$, $\gamma_{uv} = 1$, otherwise $\gamma_{uv} = O$. The labeling of a graph G=(V,E) is a bijective mapping

$$Y: V(G)+ \{1,2,\ldots,n\}. \quad ②$$

There is also a labeling of the edges

$$\lambda(e): u \rightarrow v, \quad ③$$

where $\lambda(e)$ is obviously related to the statements in equation ①. More will be said about this later.

If the weights given in equation ① are assembled into array form, the result is the adjacency matrix, which is a matrix showing the relationships between the vertices of the graph, For many applications, the weighted adjacency matrix is normalized so that each element resides in the closed interval [0,1], e.g., the definition of fuzzy relations [2] and simple graph used in fundamental mathematical analyses [8]. In this paper, we will consider normalized graphs. However, the intervals of each element of the relational matrix (relation) are not necessarily equal.

B. Directed Graphs and Trellises

A trellis T=(V,E) of rank n is a finite-directed graph G, where, as above, V=V(G) is the vertex set and E=E(G) is the edge set. The labels assigned in equation ② above specifies the depth of the vertex in the trellis. The vertex a depth zero is called the source and the vertex at depth n is tailed the sink [9]. However, because of the sectionalization of the trellis [10], any internal vertex between zero and n maybe considered as the source or sink. Therefore, it is beneficial to define the initial (init) vertex and the final (fin) vertex of the graph. By the we mean,

$$\text{init}(e) = u \text{ and } \text{fin}(e)=v. \quad ④$$

Notice that in a more generalized form, for a sequence of edges $e_1, e_2 \ldots e_L$, where $P=e_1, e_2 \ldots e_L$, with $\text{init}(e_1)= u$ and $\text{fin}(e_L)=v$, then $P:u \rightarrow v$. This shows that we may write for the i^{th} edge e, and the $(i+1)^{th}$ edge, $\text{fin}(e_i)= \text{init}(e_{i+1})$. Following [9], the number of edges leaving a vertex v is denoted by $p^+(r)$, and the number of edges entering a vertex v by p-(v), i. e.,

$$p^+(v) = |\{e:\text{init}(e)=v\}| \text{ and } p^-(v)= |\{e:\text{fin}(e)=v\}|. \quad ⑤$$

The significance of equation ⑤ is that it reinforces the fact that paths through the graph/trellis can share common vertices and can split into separate paths later. Equation ④ and the paragraph following it suggest that we can define projection maps on the trellis such that past and future sections of the trellis graph maybe identified. This says that the associated relation maybe partitioned into a past submatrix (relation) and a future submatrix (relation). We will discuss this further in the next subsection.

C. Trellis Diagram and Linear Block Codes

Considering the labeling of the edges of graph Gin equation@, we can write the labeled trellis as $T=(V, E, \lambda)$ and from [9], [11], [12] we define a fixed (n, k, d) binary linear block code,

where n is the length of the code (also the rank of the trellis), k is the number of information bits in each codeword and d is the Hamming distance, the number of bit positions that are different between any two codewords. Notice that the labels of the trellis are taken from {0,1}, with the structure of "language semiring"[9].

We say that trellis T represents a code C if the language produced by T (the concatenation of edge weights) is identical to the linear block code of C. In other words, if we associate a length - n binary (code) word with every path from source (init[e]) to sink (fin[e]) in the trellis by concatenating the edge labels on the path, and if the set of such trellis path words is identical to the set of codewords in C, then we say that T represents C. From the coding theory literature [11] [12], and others, the weight enumerator for C, that is the total "flow" of weights from source to sink will be the generating function for the weights of the codewords. The weight enumerator process supports finding the trellis with the fewest edges, and, according to [9], there is always (up to isomorphism) a unique edge -- minimal trellis that represents C. This edge-minimal trellis conclusion is a fundamental result for the results of this paper, and also leads to the following important result.

Finally, we define the i^{th} past projection mapping codeword $\underline{C} = (C_1, \ldots C_n)$, $\pi: C \rightarrow P^i$ such that $\pi_i(\underline{C}) = (C_1, \ldots C_i)$, where the kernel of π_i, i.e., the set of codewords $\underline{C}$ such that $\pi_i(\underline{C})$ is zero from the i+1 to n. This projection is called the past subcode of *C*. Likewise, a future projection $\varnothing_i: C \rightarrow F_i$ or $\varnothing_i(\underline{C}) = (C_{i+1}, \ldots C_n)$. Then the kernel of $\varnothing_i$ is the past subcode P_i as the kernel of the π_i is the future subcode F_i. Then from the well-known "rank+ nullity = dimension" theorem from linear algebra, it follows that

$$k=p^i+f_i,\ I=0,\ldots,n \text{ and } k=p_i+f^i,\ i=0,\ldots,n \qquad ⑥$$

where p^i and f^i are the dimension of the future projections sub codes and pi and f_i are the past projection subcodes. From the equivalence of the edge - minimal trellis and a code C, the projection arguments above suggests that there are edge-minimal subtrellis sections that are equivalent to the past and future subcodes.

III. Main Results

Our main results consists of the formulation of a simple systems model where the map from the input X and the output Y is the relation R of interest. The notations for the model follow [2].

Let X be the universe of discourse of input parameters and Y be the output parameter universe of discourse. The system which maps elements in *X* onto *Y is* f which has the associated relation R. This is shown in Figure 1.

X --------------------> | f,R |-----------------------> Y

Figure 1. Simple system model

For our presentation we will use a fuzzy relation R and let subsets of X and Y to be fuzzy. That is $A \in X$ and $B \in Y$ are fuzzy subsets. Then, functionally, we may write

$$f: A \rightarrow B \qquad ⑦$$

for a crisp function or if $x \in X$ is crisp

$$B=f(x) \qquad ⑧$$

for a fuzzy functional. The associated fuzzy relation has the form in equation ⑨. Below, we will further define R as a special outer product.

$$
R = \begin{array}{c} \\ x_1 \\ x_2 \\ x_i \\ x_n \end{array}
\begin{array}{c}
\begin{array}{ccccccc} y_1 & y_2 & \cdots & y_j & \cdots & y_m \end{array} \\
\begin{bmatrix}
r_{11} & r_{12} & \cdots & r_{1j} & \cdots & r_{1m} \\
r_{21} & r_{22} & \cdots & 4_{j} & \cdots & r_{2m} \\
r_{i1} & r_{i2} & \cdots & r_{ij} & \cdots & r_{im} \\
r_{n1} & r_{n2} & \cdots & r_{nj} & \cdots & r_{nm}
\end{bmatrix}
\end{array}, \qquad ⑨
$$

where $x_i \in X$ and $y_j \in B \in Y$.

From Section II, Subsections A, B, and C we know that equation ⑨ has an associated graph, trellis and a binary linear block code under appropriate normalization, Also from [2], an appropriate lambda-cut can be applied to normalized R to create the binary relations.

Our specific relations are developed from three finite length time-series of 3-dimensional (3-D) data, $(\{X_i\}, \{Y_i\}, \{Z_i\})$, i=1, 2, ... N. The relations between the i* sample and the $(i+1)^{th}$ sample for each of the time-series is formed from the outer products. These product matrix, (relations) show a three deep motion coupling or influence that propagates three samples ahead throughout these time-series. This is shown in Figure 2.

The outer products and the resulting relational matrices which couples the i^{th} vector of influence to the (i+3) vectors for coordinates x, y, and z. The x-coordinate relationship is shown in equation ⑩. Similar relationships exist for y and z but they will not be presented here.

$$
R'_i = \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} \begin{bmatrix} x_{i+1} & x_{i+2} & x_{i+3} \end{bmatrix}
$$

$$
= \begin{bmatrix} x_i x_{i+1} & x_i x_{i+2} & x_i x_{i+3} \\ y_i x_{i+1} & y_i x_{i+2} & y_i x_{i+3} \\ z_i x_{i+1} & z_i x_{i+2} & z_i x_{i+3} \end{bmatrix} \rightarrow \begin{bmatrix} r^{xx}_{i,i+1} \\ r^{yx}_{i,i+1} \\ r^{zx}_{i,i+1} \end{bmatrix} \qquad ⑩
$$

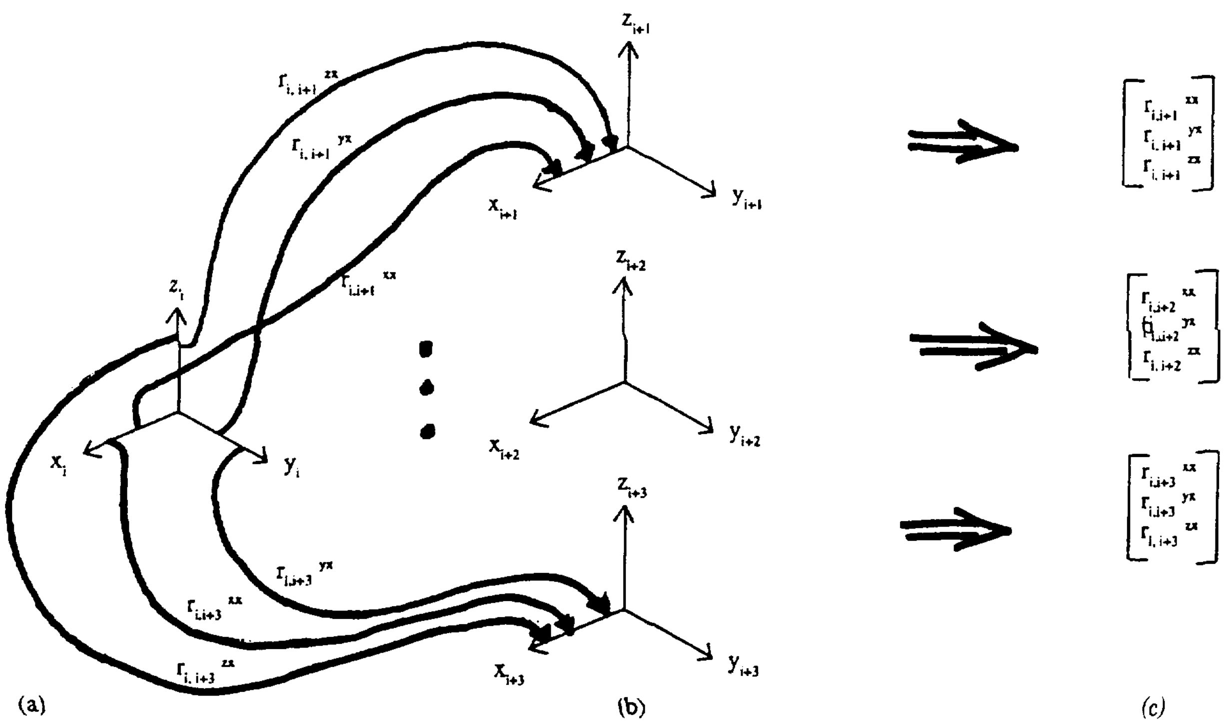


Figure 2. Influence coupling for time-series relational data

Figure 2 a) shows a typical input 3-D parameter vector coordinate. Figure 2 b) shows the depth of the three coordinates of 3-D parameters that are influenced or coupled to the parameters of the typical input vector. In Figure 2 (c), the three columns of the resulting relational matrix for the three deep x-coordinate parameters. The relation is shown in equation ⑩. The three deep influence reach of these relations represent memory of the motion or activity at i that is present at i +3. Since the time-series evolves to time N, the sequential relations represent dynamic Markov process of third-order. This Markov property will be investigated in a future paper.

From Section II A., we know that there is a weighted directed graph defined for equation (1) and the relations for y and z. Also from Section H B., these graphs can be made equivalent to edge-minimal trellises, which are in turn isomorphic to a binary linear block code. Finally, we saw in Section II C, that these edge-minimal trellises and their binary linear block codes have the sectionalization property where each codeword and trellis can be partitioned into a past section and a future section with respect to any internal vertex. Such a trellis is shown in Figure 3.

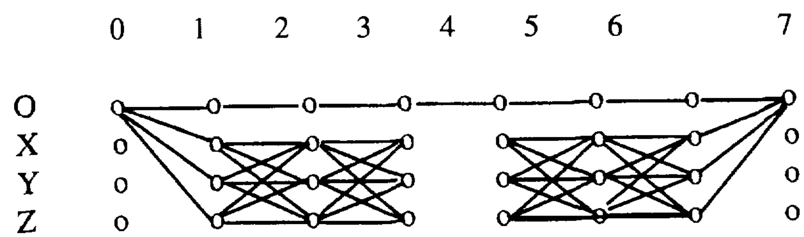


Figure 3. An edge-maximal trellis

First notice that for vertices 1-3 and 4-6, are connected by all edges. Such a trellis is called an edge-maximal trellis for states x, y, and z. Furthermore, we will take vertices 1-3 as the past, and vertices 4-6 as the future subgraph with reference to edges between vertex 3 and vertex 4.

Second, the codewords and subcodewords are defined by the binary weights (O or 1) on the edges. For example, the subcodewords of edgeweights from the source vertex 0 to the sink vertex 3 is 111 for each path. The sequences of state equivalent subcodewords are oxxx, oxxy, oxxz, ... 0000 for a total of 28 subcodewords. Obviously these are not unique codewords. However, unique codewords may be realized from real data as shown in the edge weights in Figure 4, which are subcodewords of a real time-series from a biological source are realistic edge-minimal subcodewords. This rank 3 trellis is typical of the longer sequence of codewords currently being investigated. The data is 3-D kinematics data.

	Edges		
States	0-1	1-2	2-3
0	0	0	0
x	0	0	0
y	0	1	1
z	1	0	0

Figure 4. A simple 3-bit subcode taken from real 3-D data.

Notice that a codeword or subcodeword represents a measure the time-series sojourn in each state. For Figure 4, states o and x were not visited, while state y was visited on edges e_{12} and e_{23}, and state z was visited on edge e_{01}.

For long time-series, the total sequence of codeword message will extend over the time-series and be subdivided into an integer valued number of codewords based on the past and future partition. Furthermore the actual "message" bit structure will depend on the lambda-cut (λ-cut) used on the normalized relation matrix. For a particular λ-cut, the codeword set should be unique for the given time-series. This is currently being investigated.

IV. Conclusions and Future Research

We have presented definitions of directed weighted graphs, trellises and linear block codes as ways of showing a connection between fuzzy relations and linear block codes. The connection has been shown through a special relation formed by taking the outer product of two time-series data vectors. Then the relation matrix is normalized followed by a λ-cut which results in a edge-minimal trellis and the linear binary codeword. We discuss the methods using real 3-D data.

Future plans are to investigate the Markov properties inherent in the outer product relation model and to determine the uniqueness of the linear code given the cuts chosen.

[1] S.A. Orlovski, Calculus of Decomposable properties, Fuzzy Sets, and Decisions, Allerton Press, In., NY, 1994.

[2] T. J. Ross, Fuzzy Logic with Engineering Applications, McGraw-Hill, Inc., N. Y., 1995.

[3] G.J. Klir and T.A. Folger, Fuzzy Sets, Uncertainty, and Information, Prentice Hall, Englewood Cliffs, N.J. 1988.

[4] Optimization of Fuzzy Models, W. Podrycz and J. Valente de Oliveira, IEEE Trans. on System, Man and Cybernetics - Part B: Cybernetics, Vol. 26, No. 4, August 1996.

[5] Fuzzy Sets for Intelligent Systems, Edited by D. Dubois, H. Prade, and R.R. Yager, Morgan Kaufmann Publishers, Inc., CA, 1993.

[6] Fuzzy Logic Implementation and Applications, Edited by M.J. Paytra and D.M. Mlynek, Wiley Teubner, NY, 1996.

[7] Fuzzy Mathematical Models in Engineering and Management Science, A. Kaufmann and M.M. Gupta, North-Holland, NY, 1988.

[8] M. Juvan and B. Mohar, "Laplace Eigenvalues and Bandwidth-type Invariants of Graphs," Journal of Graph Theory, Wiley -- Interscience Publication, John Wiley & Sons, In., NY, Vol. 17, no. 3, July 1993, pp. 393-407.

[9] R. J. McEliece, "On the BCJR Trellis for Linear Block Codes," IEEE Trans. Inform. Theory, Vol. 42, no. 4, July 1996.

[10] A. Lafourcode and A. Vardy, "Optimal Sectionalization of a Trellis," IEEE Trans. Inform. Theory, vol. 42, no. 3, May 1996.

[11] R. J. McEliece, The Theory of Information and Coding, Addison-Wesley, Reading, MA, 1977.

[12] E.R. Berlekamp, Algebraic Coding Theory, New York, McGraw-Hill 1968 and Laguna Hills, CA: Aegeon Park Press, 1984.

Experiments on the Effect of Counterflow on the Aeroacoustic Properties of a Supersonic Rectangular Jet

J.). M. Washington, C. Shih, F. S. Alvi and A. Krothapalli
Florida A&M University and Florida State University, Tallahassee, Florida 32316

ABSTRACT

This paper describes the results of a study on the effects of counterflow on the mixing and acoustic properties of a Mach 1.44 rectangular jet. The jet is operated over a wide range of nozzle pressure ratios encompassing design, under and over-expanded modes. As expected, the flowfield measurements indicate that counterflow enhances the jet diffusion rate by significantly increasing the mixing rates (by as much as 60°/0) of the compressible shear layers at the jet periphery. The enhanced mixing rates are observed at all nozzle pressure ratios. Acoustic measurements conducted in the anechoic facility indicate that although counterflow is very effective in eliminating screech tones at all operating conditions, it has relatively little influence in reducing the overall sound pressure -level (OASPL), which was only reduced by a few dB in the forward quadrant. It is conjectured that although screech and broadband shock-associated noise may decrease, the enhanced mixing with increasing levels of counterflow increases the turbulent mixing noise, resulting in a relatively constant OASPL.

INTRODUCTION

The low diffusion rates of supersonic jets due to the poor mixing properties of compressible shear layers have imposed a significant limitation on the design of efficient air-breathing propulsion systems which would be required to power the next generation of high-speed aircraft. Much effort has been expended over the past two decades in investigating the behavior of compressible shear layers and in pursuing ways of enhancing compressible mixing[1-6]. More recently, a unique approach has been utilized by Strykowski et al[7] in which secondary flow streams traveling in a direction opposite to the primary flow are used to generate a compressible *countercurrent shear layer*. The mixing rates of this shear layer were found to be significantly higher than the conventional coflowing compressible shear layers at convective Mach numbers (M_c) as high as two 8 However, by enhancing the mixing of compressible shear layers we have only addressed half of the problem; there is the equally important, and perhaps more intractable, problem of supersonic jet noise. To successfully design the next generation of supersonic aircraft, especially for commercial applications such as the High Speed Civil Transport (HSCT), one must also significantly reduce the jet noise generated from such vehicles.

The present study was initiated with this goal in mind. Our aim was to study the aerodynamic, and more importantly, the acoustic properties of countercurrent shear layers generated at the periphery of a supersonic rectangular jet with a design Mach number of 1.44, To better simulate realistic operating conditions, the jet was operated at ideal, over and under-expanded modes. The effect of counterflow on the mixing layer diffusion rates, screech tones, shock associated noise and mixing noise was examined in detail using pressure surveys, microphone measurements and flow visualization.

EXPERIMENTAL FACILITY AND TECHNIQUES

Test Facility

The experiments were conducted in the blow-down compressed air facility of the Fluid Mechanics Research Laboratory (FMRL). The jet flow is driven by an air compressor which is capable of supplying air at a maximum pressure of 2300 psi. The air supply consists of a bank of large storage tanks with a total capacity of 10 m', Air can also be heated by passing through an array of resistive heaters having a maximum power output of 450 kW which can increase the stagnation temperature up to 750 Kelvin. During an experiment, the stagnation pressure inside the reservoir can be maintained within ±0.5% using staged control valves,

Counterflowing Nozzle and Collar

A rectangular convergent-divergent (C-D) nozzle with a design Mach number, M_d of 1.47 (subsequent measurements indicate that the actual design Mach number is 1.44) and an aspect ratio of 4 is used for the experiment. Based on the results of previous counterflow studies[7,9], an optimized counterflow (or suction) collar was designed for the experiment, A schematic of the nozzle and collar geometry is shown in Fig. 1. To provide optical access for flow visualization, the side walls of the collar are made of acrylic plastic (plexiglass) while the upper and lower surfaces of the collar are equipped with glass windows. The small dimension (height, H) of the rectangular nozzle is 10.41 cm (0.4 1") and the gap between the nozzle and the collar surface, at the nozzle exit, is

approximately 50% of the nozzle height. The counter flowing current is produced by connecting the gaps in the suction collar to a vacuum source which consists of two regenerative vacuum pumps, connected in series.

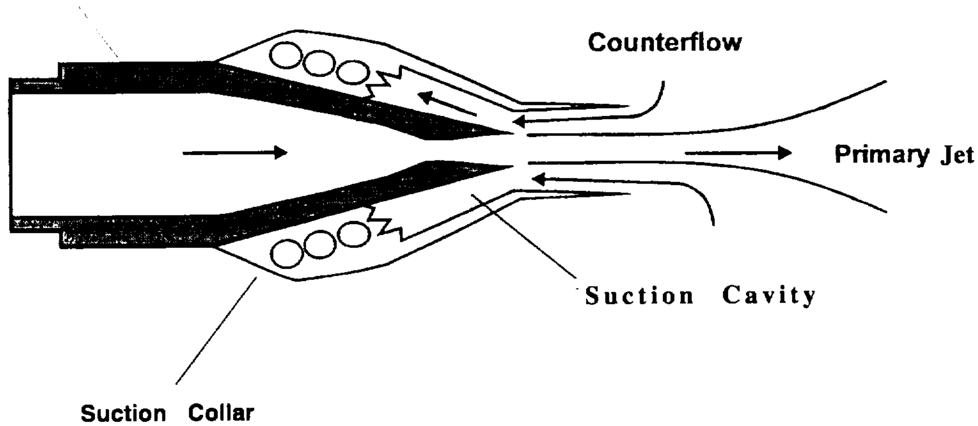


Fig. 1 - Mach 1.44 nozzle-collar geometry

Measurement Techniques

The mean aerodynamic properties of the flowfield were examined in detail via standard pitot and static pressure surveys of the jet flowfield. The pitot surveys were used to determine the extent of the potential core and the shear layer growth rates, while the static pressure profiles provide a measure of the shock-eel l structure in the jet plume. Far field acoustic measurements, emphasizing the characterization of the frequency shifts and directivity of noise radiation as a function of counterflow were conducted in the FMRL anechoic facility using a 1/4" B&K condenser microphone. (A description of the anechoic facility maybe found in Ref. 10).

Planar Laser Scattering (PLS) and white light schlieren and shadowgraphy were used to visualize the flow field. The images were recorded on a high-resolution (1048 x 1048 pixels) digital CCD camera. As the subsequent results will show, Schlieren and shadowgraph imagery is especially useful in determining the shock cell structure and for identifying the presence of acoustic radiation and screech tones emanating from the supersonic jet.

Test Conditions

As stated earlier, the primary aim of the current research is to characterize the effect of counterflow on aeroacoustic properties of the jet at design and off-design conditions. Consequently, the jet was operated over a wide range of nozzle pressure ratios, NPR, spanning a fully expanded jet Mach number (M_j) range from 1 to 2 with corresponding moderately over-expanded to moderately under-expanded flowfields. (NPR = P_0/P_e , where PO represents the total pressure and P_e represents the static pressure at the nozzle exit). A summary of the test conditions is provided in the table below. For the sake of brevity, only results from 3 selected cases, representing ideal and over and under-expanded conditions will be primarily discussed in this paper.

NPR	2.77	3.18	3.37	4.25	5.75	7.82
M_j	1.3	1.4	1.44	1.6	1.8	2.0

Table 1- Summary of test cases

RESULTS AND DISCUSSION

The effect of counterflow was examined by measuring the jet properties with and without counterflow. The amount of counterflow is measured by monitoring the pressure on the upper surface of the suction collar in the nozzle exit plane, where counterflow is roughly proportional to the vacuum pressure at that point. Hence, a higher *vacuum* pressure (i.e. lower absolute pressure) in the collar implies higher counterflow. In addition, the effect of varying the level of counterflow was also examined for each of the test cases shown in Table 1 by observing the flow fields for a minimum of two levels of counterflow -- moderate and high. Moderate counterflow corresponds to a vacuum

pressure of approximately 2.7 psi (12 psia) and high counterflow corresponds to a vacuum pressure of 4.7 psi (10 psia).

Aerodynamic Properties

Mixing Enhancement

To quantify the effect of counterflow on the potential core length, centerline mean pitot surveys were conducted for all the test cases, with and without counterflow. A summary of these measurements is shown in the plot below, where L_{core}/H is the core length non-dimensionalized by the jet height, and 'free jet' implies no counter flow.

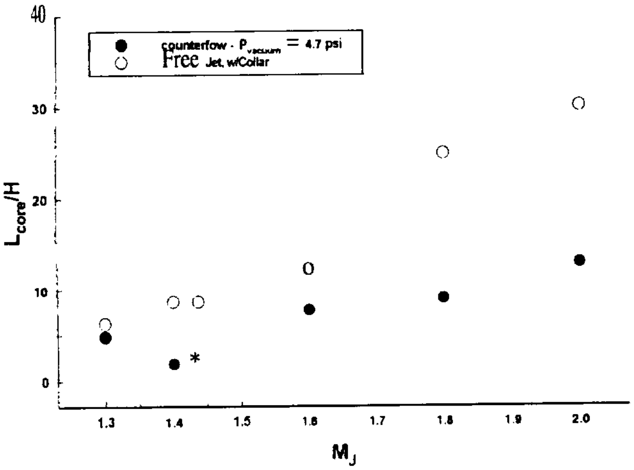
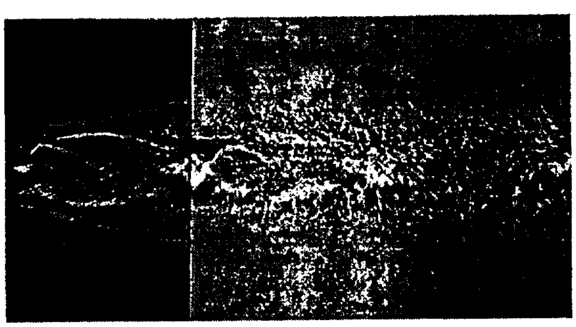


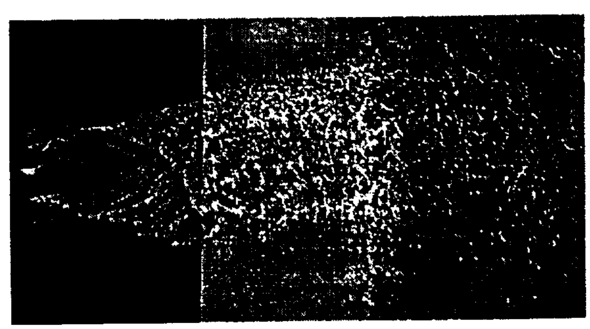
Fig. 2- Effect **of counterflow on potential core length**

As seen in the plot, even a moderate amount of counterflow significantly enhances mixing, thereby reducing the potential core length. Moreover, this mixing enhancement due to counterflow is observed at all jet operating conditions. Centerline profiles for $M_j = 1.8$ measured for various levels of counterflow indicate that there is an inverse relationship between collar vacuum pressure (degree of counterflow) and potent ial core length: for a fixed NPR, increasing counterflow enhances mixing **and** decreases the potential core extent.

Detailed pitot surveys across the shear layer were also obtained at several streamwise locations for $M_j = 1.44$ (design condition) and $M_j = 1.8$ (under-expanded). Shear layer growth rates derived from these data confirm the trends observed in Fig. 2. At a vacuum pressure, P_{vac} , of 4.7 psi, the shear layers at both Mach numbers grow at rates 70°/0 to 80% higher than the jets without counterflow. The dramatic reduction in potential core as a result of enhanced mixing due to counterflow can also be clearly seen in Fig. 3. Figure 3a shows the $M_j = 1.8$ free jet without counterflow whereas Fig. 3b depicts the same jet with counterflow at a vaccuum pressure of 5.7 psi (9 psia).



a)



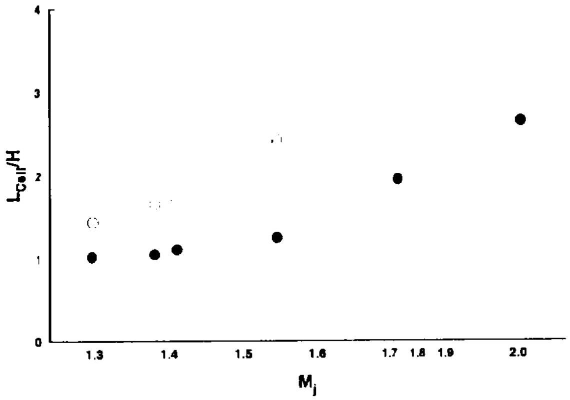
b)

Fig. 3- Effect **of counterflow on mixing for a M_j = 1.8 jet.**
a) Free Jet, no counterflow, b) Counterflow, P_{vac} = 9 psia

Shock Cell Structure

Centerline static pressure profiles were obtained for a range of NPR's or fully expanded jet numbers to examine the effect of NPR and counterflow on shock cell structure and specifically, on shock cell spacing. Several important trends are revealed from these surveys. First, as observed in the pitot surveys, the potential core lengths are dramatically reduced due to counter flow. Second, the number of shock cells in the jet plume is reduced for cases with counter flow. Third, the average shock cell spacing is reduced with increasing counterflow and lastly, the strength of the shocks decreases with increasing counter-flow. A summary of these observations is shown in Fig. 4 where Fig. 4a depicts, the average shock cell spacing (L_{cell}/H) as a function of M_j whi le Fig. 4b shows the effect of increasing the level of counterflow for a fixed M_j. Fig. 4a shows that increasing M_j increases shock cell spacing, a

trend also observed by Seiner and Norum[11] for over and under-expanded round jets. Furthermore, as depicted in Fig. 4b, for a fixed M_j, increasing counterflow (i.e. increasing vacuum pressure, P_{vac}) proportionally decreases shock cell spacing. Although not shown here, the number of shock cells also decreases linearly with increasing counter flow.



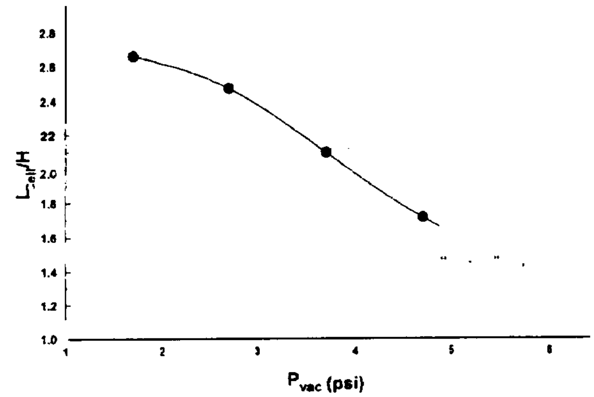


Fig. 4- Influence of counterflow on shock cell spacing. a) cell spacing vs. M_j ; b) cell spacing for M_j =1.8 as a function of counterflow

Acoustic Properties

The acoustic data presented in this section consists of far-field measurements taken in the anechoic facility. The measurements were conducted in a polar grid, at various angular locations at a fixed radius of 41.6" (R/H = 101.5). All angles and distances were measured relative to the center of the suction collar exit plane.

Screech Tones

Narrow-band noise spectra for an over-expanded case $M_j = 1.3$ are shown in Fig. 5 where the sound pressure level SPL (dB) is shown as a function of frequency. For the free jet case, screech tones and their accompanying harmonics are clearly present as manifested by the several distinct, sharp peaks. In contrast, when counterflow is present the screech tones disappear. The elimination, or at least a drastic attenuation, of screech tones due to counterflow is measured for all jet nozzle pressure ratios. Screech tones are a result of the interaction between amplified instability waves and the shock cell structure[12] ; by drastically altering the shock cell structure (see Figs. 3-5), counterflow presumably disrupts the feedback loop required for screech tone generation,

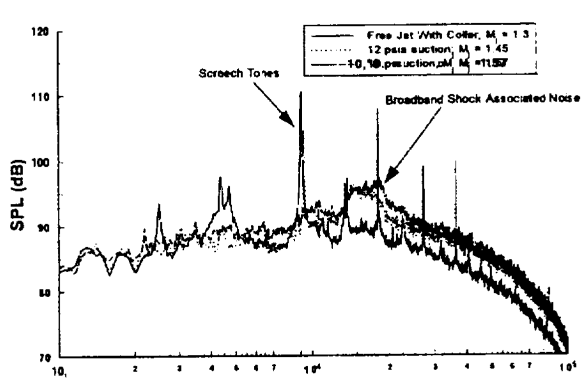


Fig. 5- Narrow band noise spectra for M_j = 1.3, with and without counter flow.

Broadband Shock-Associated Noise

While eliminating screech tones, counterflow also affects the overall spectral noise distribution, especially broadband shock-associated noise. Fig. ***6*** shows narrow-band noise spectra for M_j =1.6 for two levels of counter flow. The spectra show that peak frequencies of the broadband shock-associated noise shifts to higher frequencies with increasing counterflow (This trend can also be observed in Fig. 5, for the overexpanded jet). It is widely believed that shock associated noise is generated by the interaction between large-scale turbulent structures in the shear layer and the quasi-periodic shock cell structure[12]. According to this model, a reduction in the shock cell spacing would lead to a shift of the dominant frequencies to higher levels, consistent with the behavior observed in Fig. 6.

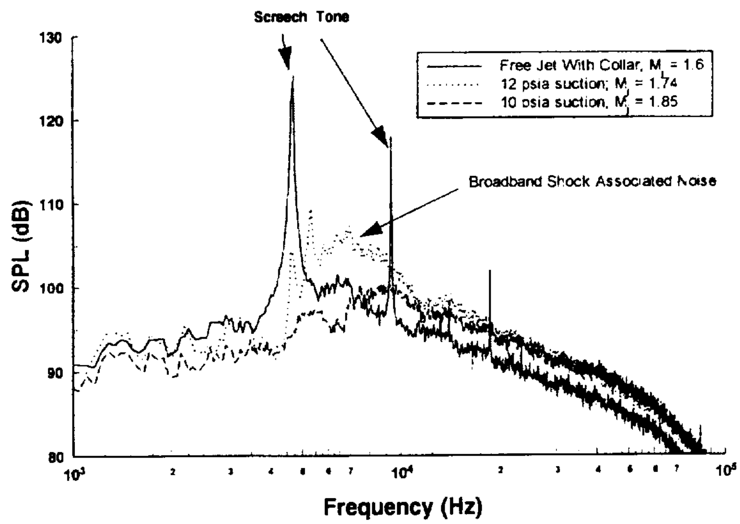


Fig. 6- Narrow band noise spectra for M_j = 1.6

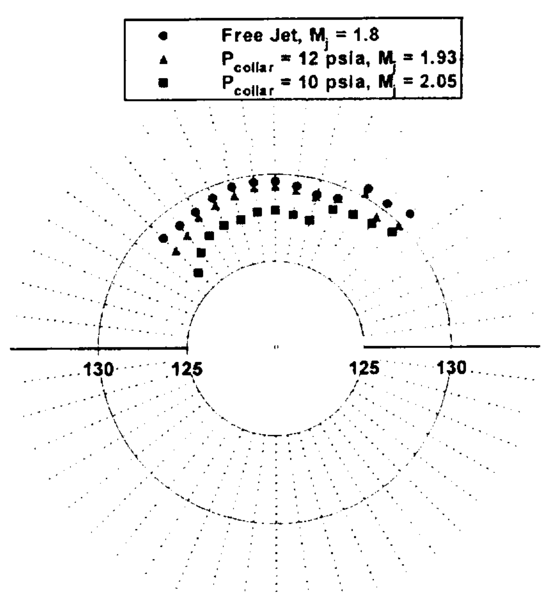


Fig. 7- Polar plot of OASPL for Mj = 1.8.

A polar plot of the overall sound pressure level (OASPL) for fully expanded jet Mach number of 1.8 is shown in Fig. 7, where each radial line represents a division of 7.5 degrees.. This plot allows us to examine the directivity as well as the overall noise levels as a function of counterflow. As seen in the figure, the OASPL decreases over the entire measurement space with the most noticeable reduction occurring in the forward quadrant. Nonetheless, even in this region, the noise reduction is fairly modest (approximately 3dB at θ = 45 deg.). Similar trends in OASPL -- as a function of counterflow -- are observed for other M_j's. There are many possible explanations for this behavior. First, a close examination of Figs. 5 and 6 reveal that although the spectral content of the noise is modified, the overall noise level is relatively unchanged. In fact, the sound pressure level at higher frequencies increases slightly which suggests that the overall sound pressure level may actually increase for some cases. Second, it is well-known that counterflowing shear layers have much higher turbulent intensities than conventional shear layers[7-9]. These highly energetic turbulent structures may not only lead to higher turbulent mixing noise, but their interaction, even with the relatively weak shock cells (in the case of counterflow), may result in higher broadband shock-associated noise levels. This would lead to an overall increase in OASPL.

CONCLUSIONS

This aeroacoustic properties of a Mach 1.44 rectangular jet operated at design, under and over-expanded modes were investigated. The aerodynamic measurements indicate that counterflow enhances the jet diffusion rate, at all nozzle pressure ratios by significantly increasing the mixing rates of the compressible jet shear layers. Acoustic measurements show that although counterflow is very effective in eliminating screech tones at all operating conditions, it has relatively little influence in reducing the overall sound pressure level (OASPL), which was only reduced by a few dB in the forward quadrant. It is conjectured that although screech and broadband shock-associated noise may decrease, the enhanced mixing with increasing levels of counterflow increases the turbulent mixing noise, resulting in a relatively constant OASPL.

REFERENCES

1. Papamoschou, D. and Roshko, A. R., "The Compressible Turbulent Shear layer; An Experimental Study," *Journal of Fluid Mechanics,* Vol. *197, 1988,* pp. 453-477.
2. Bogdanoff, D. W., "Compressibility Effects in Turbulent Shear Layers," *AIAA Journal,* Vol. 21, **No. 6**, *1983* pp. *926-927.*
3. Clemens, N. T. and Mungul, M. G., "Two-and-three-Dimensional Effects in the Supersonic Mixing Layer," *AIAA Journal,* Vol. *30, No. 4, 1992,* pp. 973-981.
4. Shau, Y. R., Dolling, D. S. and Choi, K. Y., "Organized Structure in a Compressible Turbulent Shear Layer,"*AIAA Journal,* Vol. 31, No. 8, 1993, pp. 1398-1405.
5. Naughton, J. W., "The Enhancement of Turbulent Compressible Mixing Via Streamwise Vorticity," Ph.D. Dissertation, Dept. of Mechanical Engineering, Pennsylvania State University.
6. Samimy, M., Zaman, K. B. M. Q. and Reeder. M. F., "Supersonic Jet Enhancement by Vortex Generators," A IAA Paper 91-2263, July 1991.
7. Strykowski, P. J., Krothapalli, A and Jendoubi, S., "The Effect of Counterflow on the Development of Compressible Shear Layers," *Journal of Fluid Mechanics,* Vol. 308, 1996, pp. 63-96.
8. Alvi, F. S., Krothapalli, A. and Washington, D., "Experimental Study of a Compressible Countercurrent Turbulent Shear Layer," *AIAA Journal,* Vol. 34, No. 4, 1996, pp. 728-735.
9. Strykowski, P.J. & Krothapalli, A., and Forliti, D. J. "Counterflow Thrust Vectoring of Supersonic Jets,"*AIAA Journal,* Vol. 34, No. 11, 1996, pp. 2306-2314.
10. Alvi, F. S., Krothapalli, A., Washington, D. and King, C. J., "Aeroacoustic Properties of a Supersonic Diamond-Shaped Jet," *AIAA Journal,* Vol. 34, No. 8, 1996, pp. 1562-1569.
11. Seiner, J. M and Norum, T. D. "Aerodyanmic Aspects of Shock Containing Jet Plumes", AIAA Paper **80-9065.**
12. Tam, C. K. W., "Supersonic Jet Noise", *Annu. Rev. Fluid Mech.* Vol 27, pp. 17-43, 1995.

URC97155 NASA 5154/39

NEURAL NETWORK PREDICTION OF FAILURE OF DAMAGED COMPOSITE PRESSURE VESSELS FROM STRAIN FIELD DATA ACQUIRED BY A COMPUTER VISION METHOD

Samuel S. Russell
NASA
Marshall Space Flight Center
AL 35812
(205) 544-4411

Matthew D. Lansing
University of Alabama in Huntsville
Research Institute
Huntsville, AL 35899
(205) 890-6343 ext. 281

INTRODUCTION:

This effort used a new and novel method of acquiring strains called Sub-pixel Digital Video Image Correlation (SDVIC) on impact damaged Kevlar/epoxy filament wound pressure vessels during a proof test. To predict the burst pressure, the hoop strain field distribution around the impact location from three vessels was used to train a neural network. The network was then tested on additional pressure vessels. Several variations on the network were tried. The best results were obtained using a single hidden layer.

SDVIC is a fill-field non-contact computer vision technique which provides in-plane deformation and strain data over a load differential. This method was used to determine hoop and axial displacements, hoop and axial linear strains, the in-plane shear strains and rotations in the regions surrounding impact sites in filament wound pressure vessels (FWPV) during proof loading by internal pressurization. The relationship between these deformation measurement values and the remaining life of the pressure vessels, however, requires a complex theoretical model or numerical simulation. Both of these techniques are time consuming and complicated. Previous results using neural network methods had been successful in predicting the burst pressure for graphite/epoxy pressure vessels based upon acoustic emission (AE) measurements in similar tests (Walker, J. L., Hill, E. v. K., Workman, G. L., Russell, S. S., "A Neural Network/Acoustic Emission Analysis of Impact Damaged Graphite/Epoxy Pressure Vessels," *American Society for Nondestructive Testing, 1995 Spring National Conference, 20-24* March 1995). *The* neural network associates the character of the AE amplitude distribution, which depends upon the extent of impact damage, with the burst pressure. Similarly, higher amounts of impact damage are theorized to cause a higher amount of strain concentration in the damage effected zone at a given pressure and result in lower burst pressures. This relationship suggests that a neural network might be able to find an empirical relationship between the SDVIC strain field data and the burst pressure, analogous to the AE method, with greater speed and simplicity than theoretical or finite element modeling.

The process of testing SDVIC neural network analysis and some encouraging preliminary results are presented in this paper. Details are given concerning the processing of SDVIC output data such that it may be used as back propagation neural network (BPNN) input data. The software written to perform this processing and the BPNN algorithm are also discussed. It will be shown that, with limited training, test results indicate an average error in burst pressure prediction of approximately six percent,

SPECIMEN:

This study is part of a larger damage assessment program concerning impact damaged FWPV. The specimen utilized for this study conform to ASTM standard D2585-68 (1985), and represent sub-scale simulated solid rocket motor casings or fuel storage vessels with a 14.6 cm diameter. The bottles were formed by a series of helical and hoop plies, with the outermost being a hoop, as shown in Figure 1. At the NASA Marshall Space Flight Center's Productivity Enhancement Complex, DuPont Kevlar fibers were wet wound with Dow DPL862/W epoxy resin around an inner rubber bladder on a sand mandrel and then rotisserie cured to form each specimen. The undamaged burst pressure was approximately 20.7 MPa.

APPARATUS AND PROCEDURE:

An air driven water pump was used to internally pressurize each specimen as shown in Figure 2. This figure illustrates the SDVIC data acquisition hardware. Approximately 6.45 cm^2 around the impact zone on each impacted bottle, or at a random location on each un-impacted bottle, was viewed by a CCD camera with illumination by two 500 watt halogen quartz shop lamps. A random black and white speckle pattern was applied to the region of interest on each bottle by over-spray from ordinary flat or low gloss spray paint to assist in the image correlation process. A PC based image digitization board was used to acquire images during holds in the pressurization cycle. These images correspond to approximately 0,8,

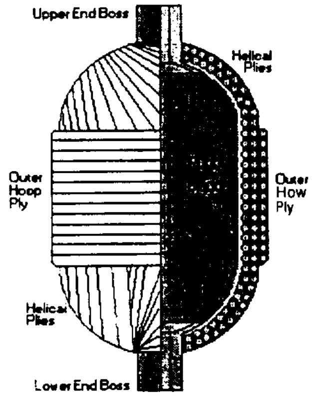


Figure 1. Specimen Geometry

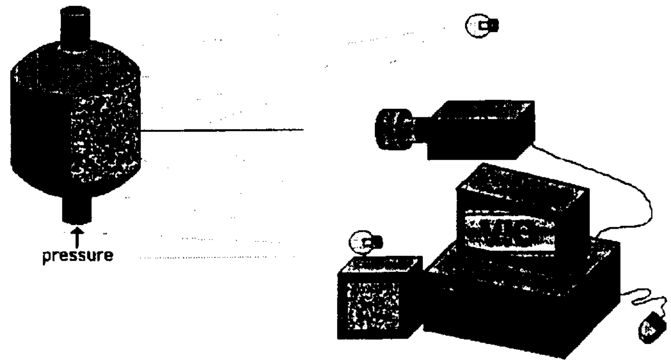


Figure 2. Data Acquisition Apparatus

17, 25, and 33 percent of the undamaged burst pressure for an un-impacted bottle. To minimize vibration effects, five frames were averaged for each image.

The SDVIC image processing software was used to correlate each non-zero image with the image acquired at O pressure. This software utilizes a pattern recognition algorithm to determine with sub-pixel resolution the relative position, and thus deformation, of small image subsets between two images. An automated routine repeats this process for a grid of subsets covering the entire region of interest, resulting in tabulated and false color plotted full-field displacement and strain data. Figure 3 and 4 are a false color plot of the hoop strains and a photograph of an impact damaged graphite/epoxy pressure vessel respectively. The underlying hoop plies had ripped vertically as a result of the impact an a previous low level pressurization. The SDVIC software package is available from NASA COSMIC.

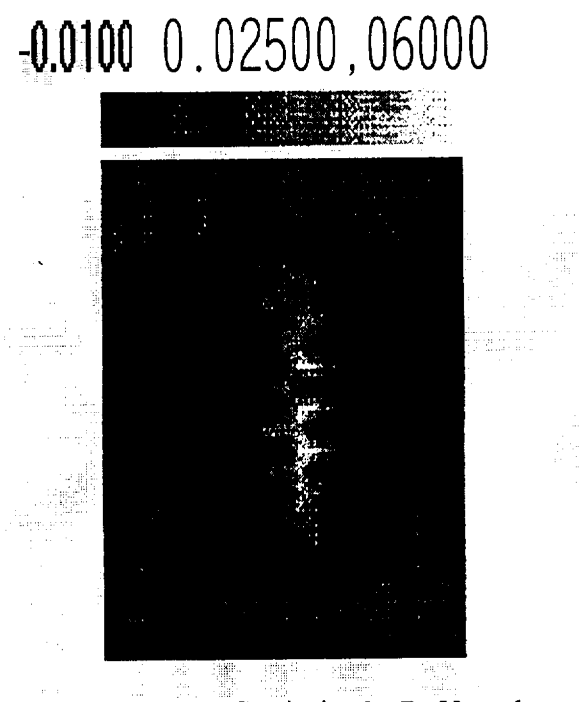


Figure 3. Hoop Strain in Gr-Ep Vessel.

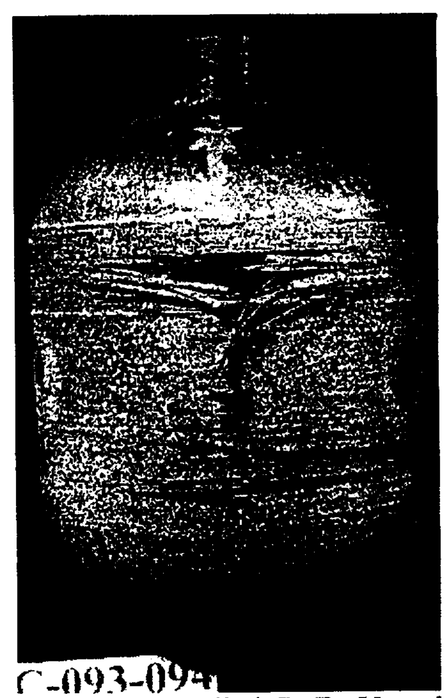


Figure 4. Failed Gr-Ep Vessel.

DATA REDUCTION:

In the case of AE neural network analysis, the amplitude distribution is divided into a series of discrete categories, with the population of each supplied as the input to the corresponding input layer neuron. The six deformation fields (in-plane displacement and the inplane strains and rotations) generated by SDVIC processing are each composed from a grid of 134 x 134 measurements, for a total of 17,956 data points. It is not practical to attempt processing with this many input neurons. Therefor, the tabulated SDVIC deformation field data is summarized by a group of distributions analogous to the AE amplitude distribution, With some qualitative foresight as to the failure characteristics of these specimens, the hoop strain was chosen as the single parameter which most strongly represents the damage, and is the only input to the neural net.

It is theorized that an increase in impact damage severity, corresponding to a decrease in burst pressure, will cause more of the strain field to contain higher strain values due to the strain concentrating effect of that damage. That is, at a given internal pressure an un-impacted bottle maybe expected to have a narrow distribution of hoop strain values about some average. An impacted bottle with the pressure, field of view, and all image correlation parameters repeated should have a lower and wider distribution due to a shift toward higher strain values. Figure 5. illustrates that a bottle with a burst pressure of 17.9 MPa has a taller, narrower distribution than one with a burst pressure of only 11.7 MPa. From another point of view, the area under this curve represents a strain x area product, which resembles a strain energy.

The minimum and maximum strain varies from specimen to specimen, and the neural network approach requires the same number of inputs for each specimen. Thus, each individual strain field was converted to a strain distribution with the same number of categories, but not necessarily the same category ranges. For example, if there are 20 strain categories and input neurons, then the first input neuron always receives the number of data points which fall in the lowest twentieth (or five percent) of the strain distribution.

The hoop strain distributions used in this test are shown in Figure 6. in order of increasing burst pressure from front to back. This set of nine specimen represent a range of impact damage levels from none to that which reduced the actual burst pressure by approximately one quarter of the undamaged value. Seventeen strain categories and input neurons were used. The data used here was obtained during a 6.89 MPa hold in the proof cycle, or at approximately one third of the undamaged burst pressure. This level of pressurization should cause no damage to an un-impacted specimen.

NEURAL NETWORK PROCESSING:

A software program called VICNet was written to convert the tabulated SDVIC strain field data into strain distribution tables for the entire population of specimens at once. The program then input the strain distributions and actual burst pressures for three of the nine bottles, which had been designated as the training data set. The VICNet back propagation neural network (BPNN) routine analyzed the training set to adjust internal weights and biases such that the neural network output was within an average of 5°/0 deviation from each actual burst pressure. In addition to the seventeen input neurons and single output neuron, a hidden layer of 3 intermediate neurons were used. The results of this training, which required only 85 iterations and less than 10 seconds, are shown in Table 1.

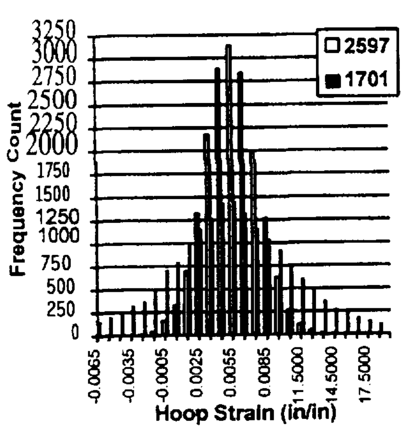


Figure 5. Comparison of Strain Distributions for Different Burst Pressures

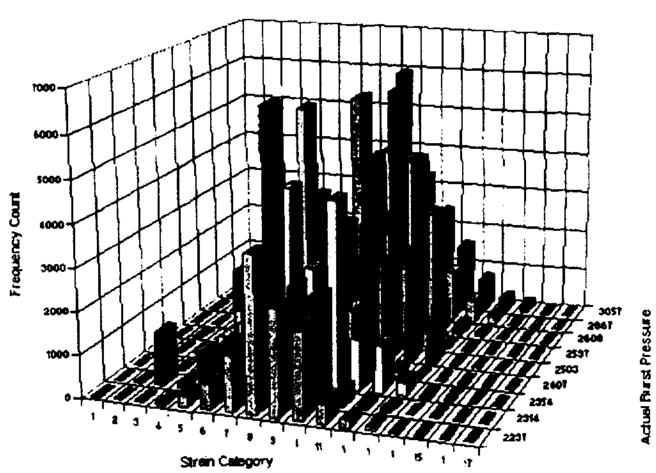


Figure 6. VICNet Hoop Strain Distributions for BPNN Input

Table 1. Training Results		
Actual Burst Pressure (MPa)	VICNet-BPNN output	% Error
15.4	15.9	3.4
18.0	18.7	4.3
21.1	19.5	-7.3
Average Magnitude Difference		5.0

Table 2. Test Results		
Actual Burst Pressure (MPa)	VICNet-BPNN output	% Error
15.9	15.1	-5.2
16.2	16.5	2.0
17.2	18.0	4.5
17.9	19.2	7.2
19.7	18.2	-8.0
16.5	15.2	-8.2
Average Magnitude Error		5.9

The VICNet BPNN routine then input the strain distributions for the remaining specimens which were designated as the test data set. These distributions were processed by the neural network in a single pass using the internal weights and biases determined from training, The actual burst pressures were in no way accessible to the software algorithm. The test results are shown in Table 2. Multiple independent repetitions of this training and testing have yielded similar average uncertainties.

CONCLUSIONS:

It has been shown that a back propagation neural network routine can, with some degree of accuracy, be trained to predict the burst pressure of impact damaged structures based upon SDVIC strain field data collected from other similar structures. In the case of the filament wound pressure vessels studied here, an average testing error of about six percent has been demonstrated. To be completely thorough, future testing will be conducted in which this method is used to predict the burst pressure of the specimen prior to actual burst testing. It will also be determined whether similar analysis at lower proof loads will provide similar levels of uncertainty. Further research will also be conducted to determine the extent to which a neural network which has been trained on SDVIC data from one size of specimen may be used to predict failure of another size of similar specimen. If this is successful, then a network which has been trained on an appropriate set of smaller, less expensive specimen may be used to predict the failure of a larger, more expensive service article which may have sustained some form of damage.

URC97156

A REPRESENTATION THEOREM FOR LINEAR SYSTEMS

5,55/67

Irwin W. Sandberg
Department of Electrical and Computer Engineering
The University of Texas at Austin
Austin, Texas 78712-1084
e-mail: sandberg@uts.cc.utexas.edu

Abstract

The cornerstone of the theory of discrete-time single-input single-output linear systems is the idea that every such system has an input-output map H that can be represented by a convolution or the familiar generalization of a convolution. This thinking involves an oversight that was recently corrected by adding an additional term to the representation, Here we give a corresponding result for the important case of representations of causal continuous-time system maps that take bounded functions into bounded functions.

I. INTRODUCTION

The cornerstone of the theory of discrete-time single-input single-output linear systems is the idea that every such system has an input-output map H that can be represented by an expression of the form

$$(Hx)(n) = \sum_{p=-\infty}^{\infty} h(n,p)x(p) \tag{1}$$

in which x is the input and h is the system function associated with H in a certain familiar way. It is widely known that this, and a corresponding representation for time-invariant systems in which $h(n,p)$ is replaced with $h(n-p)$, are discussed in many books. Almost always it is emphasized that these representation; hold *for all* linear input-output maps H. On the other hand, in [1, p. 98] attention is directed to material in [2, p. 58] which shows that certain time-invariant H's in fact do not have convolution representations. [1] This writer does not claim that these H's *are* necessarily of importance in applications, but he does feel that their existence shows that the analytical ideas in the books are flawed.[2]

In [3] it shown that, under some mild conditions concerning the set of inputs and H, (1) becomes correct if an additional term is added to the right side. More specifically, it is shown that

$$(Hz)(n) = \sum_{p=-\infty}^{\infty} h(n,p)x(p) + \lim_{k\to\infty}(HE_k x)(n)$$

[1]The claim in [1] that there exists a discrete-time time-invariant causal H that has no convolution representation is correct, but it may not be clear that the argument given there actually shows this. Specifically, it may not be clear from what is said in [1] that the pertinent map constructed there is causal and time-invariant. However, it is not difficult to modify what is said so that it establishes the claim (see [2, p. 58] or the proof of the proposition in Section G.2 of the appendix of [3]),

[2]The oversight in the books is due to the lack of validity of the interchange of the order of performing a certain infinite sum and then applying (I-I.)(n). The infinite sum at issue clearly converges pointwise, but that is not enough to justify the interchange.

for each n, in which h has the same meaning as in (1), and $E_k x$ denotes the function given by $(E_k x)(p) = x(p)$ for $|p| > k$ and $(E_k x)(p) = 0$ otherwise. This holds whenever the input set is the set of bounded functions, the outputs are bounded, and H is continuous. In particular, we see that in this setting, an H has a representation of the form given by (1) if and only if

$$\lim_{k \to \infty} (H E_k x)(n) = O$$

for all x and n. Since this is typically a very reasonable condition for a system map H to satisfy, it is clear that the H's that cannot be represented using just (1) are rather special.

A natural question that arises is whether a similar need for an additional term arises in connection with the theory of representations of continuous-time system maps. Here we show that it does. More specifically, we consider causal input-output maps H that take bounded functions into bounded functions. We show in the next section that if H is continuous and satisfies a certain condition A. 1 that is often met, then

$$(\mathbf{Hz})(\mathbf{t}) = \int_{-\infty}^{t} h(t, \tau) x(\tau)\, d\tau + \lim_{a \to -\infty} (H P_a x)(t)$$

for all t, in which h has the usual impulse-response interpretation and $(P_a x)(\tau) = x(\tau)$ for $\tau \leq a$ and $(P_a x\ \tau) = O$ otherwise. In the appendix an example is given of an H for which $\lim_{a \to -\infty} (H P_a x)(t)$ is not always zero.

IL REPRESENTATION RESULT

2.1. Preliminaries

Let F be either the set $\mathbb{R}$ of real numbers or the set of complex numbers, and let L_∞ denote the normed linear space of bounded F-valued Lebesgue measurable functions x defined on $\mathbb{R}$, over the field F, with the norm $\| .11$ given by $\|x\| = \sup_{t \in \mathbb{R}} |x(t)|$. We use L_1 to denote the set of integrable F-valued functions on $\mathbb{R}$.

For each $a \in \mathbb{R}$ define maps T_a, Q_a, and P_a from L_∞ into itself by $(T_a x)(t) = x(t - a)$, $(Q_a x)(t) = x(t)$, $t > a$ and $(Q_a x)(t) = O$ otherwise, and $(P_a x)(t) = x(t), t \leq a$ and $(P_a x)(t) = O$ otherwise. For each positive σ let w_σ denote the element of L_∞ given by Win(t) = σ^{-1} for $t \in [0, \sigma)$ and zero otherwise.

Throughout Section H H denotes a linear map of L_∞ into itself. We say that *His* causal and *time* invariant, respectively, if (as usual) $P_a H = P_a H P_a$ for each a and $T_a H = H T_a$ for each a. In the next section, we refer to the following two conditions.

A. 1. For each t and a in $\mathbb{R}$ with a < t, there is a real constant $c_{t,a}$ such that

$$|(Hu)(t)| \leq c_{t,a} \int_a^t |u(\tau)|\, d\tau$$

for u ● L_∞ with $u(t) = O$ for $t < a$.

A.2. $\sup_a |(H Q_a u)(t)| < \infty$ **for** each $u \in L_\infty$ and each $t \in \mathbb{R}$.

Condition A.2 is met whenever H is continuous because then $|(HQ_a u)(t)| \leq \|HQ_a u\| \leq \|H\| \cdot \|Q_a u\| \leq \|H\| \cdot \|u\|$. Condition A.1 is also often satisfied.

In the next section, when we say that a limit exists we mean that it exists as a number in F.

2.2. Our Theorem

The following is our main result.

Theorem Let H be causal and such that A.1 and A.2 are met. Then the following holds for each $t \in \mathbb{R}$ and each $x \in L_\infty$.

(i) $h(t,\tau) := \lim_{\sigma \to 0}(HT_\tau w_\sigma)(t)$ exists for almost all $\tau \in \mathbb{R}$, $h(t,\cdot)$ is Lebesgue measurable on $\mathbb{R}$, $h(t,\cdot)$ is essentially bounded on each finite interval in $\mathbb{R}$, and $h(t,\cdot) \in L_1$.

(ii) $\lim_{a \to -\infty}(HP_a x)(t)$ exists.

(iii) We have

$$(Hx)(t) = \int_{-\infty}^{t} h(t,\tau)x(\tau)\,d\tau + \lim_{a \to -\infty}(HP_a x)(t).$$

Proof

Let $t \in \mathbb{R}$ and $x \in L_\infty$ be given. Using A1, by a direct modification of material in the proof of Theorems 1 and 2 of [4], and with $a \in (-\infty, t)$ arbitrary, $h(t,\tau) := \lim_{\sigma \to 0}(HT_\tau w_\sigma)(t)$ exists for almost all $\tau \in (a, t)$, $h(t,\cdot)$ is Lebesgue measurable and essentially bounded on (a, t), and we have

$$(Hu)(t) = \int_a^t h(t,\tau)u(\tau)\,d\tau + (HP_a u)(t). \tag{2}$$

for each $u \in L_\infty$.

In particular, $h(t,\tau)$ exists for almost all $\tau \in (-\infty, t)$, $h(t,\cdot)$ is Lebesgue measurable on $(-\infty, t)$, and with u given by $u(\tau) = h(t,\tau)^*/|h(t,\tau)|$ if $h(t,\tau) \neq 0$ [$h(t,\tau)^*$ is the complex conjugate of $h(t,\tau)$] and $u(\tau) = 0$ otherwise for almost all $\tau \in (-\infty, t)$,

$$\int_a^t |h(t,\tau)|\,d\tau = H(Q_a u)(t)$$

for each a. Thus, by A.2 and the observation that $h(t,\tau) = 0$ for $\tau > t$, we have (i). By (2) with $u = x$, since y defined by $y(t) = h(t,\tau)x(\tau)$ for $t \in \mathbb{R}$ belongs to L_1,

$$\lim_{a \to -\infty} \int_a^t h(t,\tau)x(\tau)\,d\tau = \int_{-\infty}^t h(t,\tau)x(\tau)\,d\tau$$

and we have (ii) and (iii). This completes the proof.

For H time invariant the conclusion of the theorem holds without the hypothesis that A.2 is satisfied. This is true because A.2 is used only to establish that $h(t,\cdot)$ belongs to L_1, and for H time invariant $h(t,\cdot) \in L_1$ follows from (2) which implies that z given by

$$z(t) = \int_a^t h(t-\tau)u(\tau)\,d\tau,\ t > a$$

equals $(HQ_a u)(\cdot)$ restricted to $t > a$ which is bounded for each $u \in L_\infty$. Here we have written $h(t-\tau)$ for $h(t,\tau)$.

An example of a bounded linear causal map from L_∞ into itself that does not meet A. 1 is the identity map I, but $\lim_{\sigma \to 0}(IT_\tau w_\sigma)(t)$ vanishes for almost every τ and $\lim_{a \to -\infty}(IP_a x)(t) = 0$ for every t and x. This shows that the theorem becomes false if A. 1 is omitted.

In the appendix an extension result for time-invariant maps defined on L_∞ is given from which it follows that there are maps H of the kind addressed by the theorem for which the additional term is not always zero.

A result for noncausal input-output maps along the lines of our theorem can be proved in a similar way. In this case the extra term involves also the behavior of the input for large values of its argument, as it does in the discrete-time case discussed in Section I. And corresponding results in the setting of inputs and outputs of a finite number of variables can be proved starting with the approach described in [5].

111. APPENDIX: AN EXTENSION PROPOSITION

We begin with some additional preliminaries: Let M denote a linear manifold in L_∞ that is closed under translation in the sense that $T_aM = M$ for each a, where T_a is as defined in Section 2.1 except that here the domain of T_a is M. Assume also that $y \in M$ implies that z $\in M$, where $z(t) = y(t)$ for $t \leq O$, and $z(t) = O$ otherwise. We do not rule out the possibility that $M = L_\infty$.

Let A be a linear map of M into L_∞. Such an A is *time-invariant* or *causal*, respectively, if A is time-invariant or causal in the sense of Section 2.1 (with obvious modifications to accommodate the more general domain M). It is *bounded* if $\|A\|_M := \sup\{ \|Ax\|: x \in M, \|x\| \leq 1\} < \infty$, in which $\|.\|$ is the norm in L_∞. Our result is the following.

Proposition: Let A be shift invariant and bounded. Then there exists a bounded linear shift-invariant map B from L_∞ into itself that extends A in the sense that B is causal if A is causal and $Bx = Ax, x \in M$.

Proof

By the shift invariance of A, *we* have $(Ax)(t) = (AT_{-t}x)(0)$ for all t and all $x \in M$. The map $(A.\ )(0)$ is a bounded linear functional on M, because

$$|(Ay)(0)| = |(AT_{-a}T_ay)(0)| = |(AT_ay)(a)| \leq \sup_t |(AT_ay)(t)| \leq \|A\|_M \cdot \|T_ay\| = \|A\|_M \cdot \|y\|$$

for $y \in M$. When A is causal, $(A.\ )(0)$ has the property that $(Ay)(0) = O$ for any $y \in M$ for which $y(t) = O$ for $t \leq 0$. By the Hahn-Banach theorem [6, pp. 178 and 181] there is a bounded linear functional $\mathcal{F}$ that extends $(A\cdot)(0)$ to all of L_∞. Set $\mathcal{G} = \mathcal{F}$ if A is not causal, and if A is causal define $\mathcal{G}$ on L_∞ by $\mathcal{G}y = \mathcal{F}P_0y$. Define B on L_∞ by $(Bx)(t) = \mathcal{G}T_{-t}\,x$. It is easy to check that B is a linear shift-invariant bounded map into L_∞, that B extends A to L_∞, and that B is causal if A is causal.[3] This completes the proof.[4]

Since the set L of elements x of L_∞ such that $x(t)$ approaches a limit as $t \to -\infty$ is a linear manifold that is closed under translation, and since

$$(Ax)(t) = \lim_{\tau \to -\infty} x(\tau)$$

defines a shift-invariant bounded causal linear map of L into L_∞, it follows from our proposition that there exist causal time-invariant maps H of the kind addressed by our theorem for which the term $\lim_{a\to-\infty}(HP_ax)(t)$ is not always zero. More explicitly, the associate B via our proposition of the A just described meets A. 1 and satisfies $\lim_{a\to-\infty}(HP_ax)(t) = \lim_{\tau\to-\infty} x(\tau)$ for x $\in L$. And an example of an H of the type addressed by the theorem for which the additional term is not always zero and H is not shift invariant is obtained by adding to this B any linear bounded causal map of L_∞ into itself that meets A. 1, is not shift invariant, and has a representation without an additional term.

[3]It is also true that B can be chosen so that it preserves the norm of A, *in* the sense that $\|B\| = \|A\|_M$.

[4]A proposition similar to the one above, but for discrete-time maps, is given in [3].

REFERENCES

[1] S. P. Boyd, "Volterra Series: Engineering Fundamentals," dissertation, Univ. California, Berkeley, 1985.

[2] L. V. Kantorovich and G. P. Akilov, *Functional Analysis*, Oxford: Pergamon, 1982.

[3] I. W. Sandberg, "Multidimensional Nonlinear Myopic Maps, Volterra Series, and Uniform Neural-Network Approximations," in D. Docampo, A. Figueiras-Vidal, and F. Perez-Gonzalez (eds.), *Intelligent Methods in Signal Processing and Communications*, Birkhauser Boston, 1997 (selected papers from the Fourth Bayona Workshop on Intelligent Methods for Signal Processing and Communications, Bayona, Spain, June 1996),

[4] I. W. Sandberg, "Linear Maps and Impulse Responses ," *IEEE Transactions on Circuits and Systems*, vol. 35, no. 2, pp. 201-206, February 1988.

[5] D. Ball and I. W. Sandberg, "g- and *h*- Representations for Nonlinear Maps," *J. Mathematical Analysis and Applications*, Vol. 149, No. 2, July 1990.

[6] G. Bachman and L. Narici, *Functional Analysis, New* York: Academic Press, 1966.

Page intentionally left blank

 N11:17

Microstructural and Mechanical Properties Evaluations of Titanium Foils Processed via the Melt Overflow Process

5156/26

M.L. Weaver* and H. Garmestani[1,2]

[1]Center for Nonlinear& Nonequilibrium Aeroscience
Florida A&M University
1800 E. Paul Dirac Rd.
Tallahassee, FL 32306-4005

[2]Department of Mechanical Engineering
FAMU-FSU College of Engineering
2525 Pottsdamer St.
Tallahassee, FL 32310-2175

Introduction

The processing of titanium foils by conventional ingot metallurgy (IM) techniques involves casting ingots, hot forging into billets, followed by several hot rolling, heat treatment, and surface grinding sequences to produce plate or strip that is suitable for cold rolling to foil gauge. Using these techniques, processing losses exceeding 50% are not uncommon making commercial production of titanium foils very expensive. Recently, Gaspar *et* al. [1-3] have reported success in direct casting sheets of conventional titanium alloys and titanium-based ordered intermetallic compounds using a single-chill-roll casting technique called melt overflow rapid solidification technology (MORST). Using this technique, near-net-shape foils have been continuously cast using into -500 μm thick x 10 cm wide x 3 m long sheets which were successfully ground, cold rolled, or hot pack rolled to foil gauge (≤ 100 μm thickness). In comparison to IM processing techniques, the potential advantages of foil production from direct cast (DC) strips are improved purity, increased chemical homogeneity, and a reduction in processing losses resulting in lower processing costs. While the microstructures, mechanical properties, and textures of IM titanium alloys have been extensively characterized, they have not yet been addressed for DC titanium. In the present work, the microstructure, mechanical properties, and crystallographic textures developed in DC strips and in cold rolled foils produced from DC strips are compared with those of IM titanium foils.

Experimental Procedure

Titanium strips were cast in the plasma melt overflow furnace at Ribbon Technology Incorporated, Columbus, Ohio. The plasma melt overflow furnace combines plasma arc melting in a cold copper hearth with MORST by rotating the cold copper hearth about the same axis of rotation as the chill roll to overflow liquid onto the circumference of the chill roll [1-3]. The chemical compositions of the DC titanium strip/foil and of a conventional ingot metallurgy (IM) foil supplied for comparison purposes are given in Table 1. The DC material was prepared from conventional purity (CP) titanium [2]. However, during processing of the DC strip used in this investigation, approximately 20% Ti-6Al-4V scrap was accidentally mixed with the CP-Ti scrap. This addition was not discovered until the alloy strip had been cast, cold rolled and annealed (CR).

Portions of the DC strip were supplied to Texas Instruments, Materials and Controls Division, Attleboro, MA for cold rolling. Cold rolling was accomplished in two rolling and annealing steps where the strip was initially cold rolled to approximately 50% of it's original

thickness followed by vacuum annealing at 700°C for two hours. The strip was then rolled to a final thickness of 0.17 mm and again vacuum annealed at 700°C for two hours. The IM sheet (O. 15 mm thickness) was processed using conventional techniques.

Table 1. Chemical compositions of titanium strips/foils investigated

Processing Method	wt. %				Wt. ppm			
	Ti	Al	v	Fe	o	N	H	c
DC	balance	1.29	0.93	0.19	3600	450	8	150
IM	balance	0.043	0.008	0.12	1960	130	57	380

DC= direct cast, IM = ingot metallurgy

For texture analysis, portions of each alloy strip were cut into small pieces, mechanically polished, and etched with Krolls reagent to remove arty residual deformation layers. Texture variations were measured using the x-ray diffraction technique on a Philips X'Pert PW3040 MRD x-ray diffractometer operating at 40 kV and 45 mA. The following incomplete pole figures were measured using Ni filtered CuKα radiation to determine textures in the α phase: {0002}, $\{1\ 0\bar{1}0\}, \{10\bar{1}1\}, \{11\bar{2}0\}$, and $\{11\bar{2}2\}$. The pole figure data was analyzed using the popLA software package [4].

Dog-bone shaped tensile specimens 45 mm long with a gage section 11.4 mm x 6.4 mm were machined from the cold rolled foils, Multiple specimens were machined parallel to and perpendicular to the rolling direction. Tensile tests were performed at room temperature on a computer interactive ATS Model 1630 universal testing machine operated at constant crosshead velocities corresponding to an initial strain rate of 7.4×10^{-3} S-l.

Results and Discussion

Microstructure

In the DC strip, equiaxed grains with grain diameters approaching 35 µm were observed. During solid-state cooling, the alloy transformed martensitically to a mixture of acicular α and retained β phases. A typical optical microstructure is shown in Figure 1. In general, the strip was fully dense and contained no visible porosity or cracks. After cold rolling and annealing, microstructure consisting of fine equiaxed α grains were observed in both the DC and IM foils. Nominal grain diameters were 18 µm in the DC foil and 14 µm in the IM foil (Figure 2).

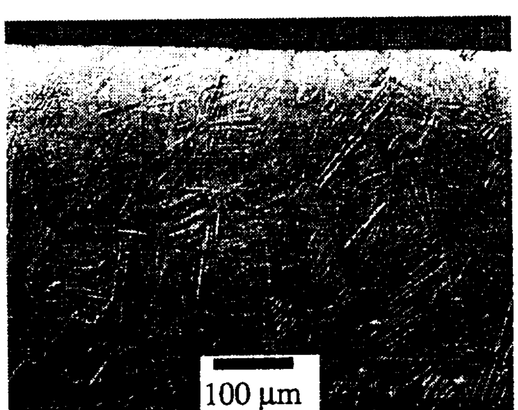


Figure 1. Optical micrograph of DC titanium strip. Casting direction is horizontal.

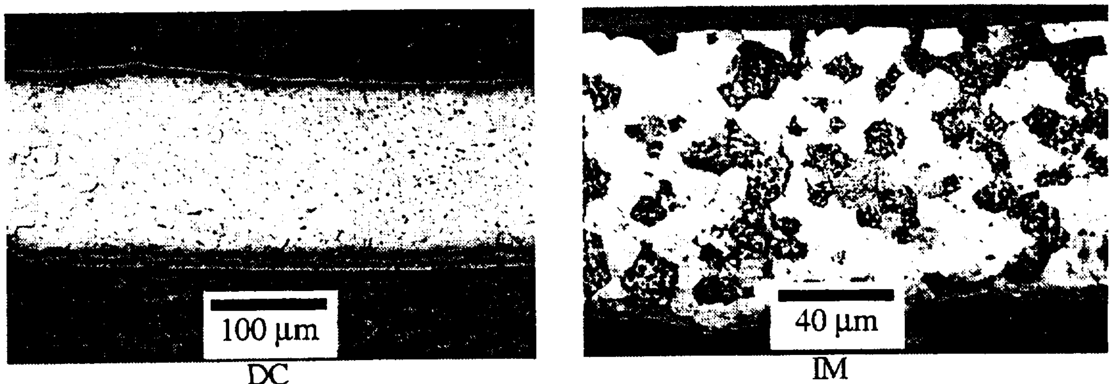


Figure 2. Optical micrographs of DC and IM titanium foils used in this study, The rolling direction is horizontal

Texture

Texture can be described as a non-random distribution of grain orientations that occurs during the manufacture of materials. Significantly different textures can result from solidification, deformation, recrystallization, and phase transformations and may lead to anisotropic mechanical properties. The experimental textures for the DC strip and for the DC+CR and IM+CR foils are represented in the pole figures displayed in Figures 3 and 4. In the DC strip, relatively weak (-1.8 x random) fiber textures were observed with the major poles oriented parallel to the strip normal in the { 11 $\bar{2}$0 } direction with some components nearly parallel to the casting and transverse directions in the {0002} pole figure. Textures where the basal plane is parallel to either the chill or wheel surface of the specimen are typical in rapidly solidified hcp metals produced via conventional chill casting or melt spinning [5,6].

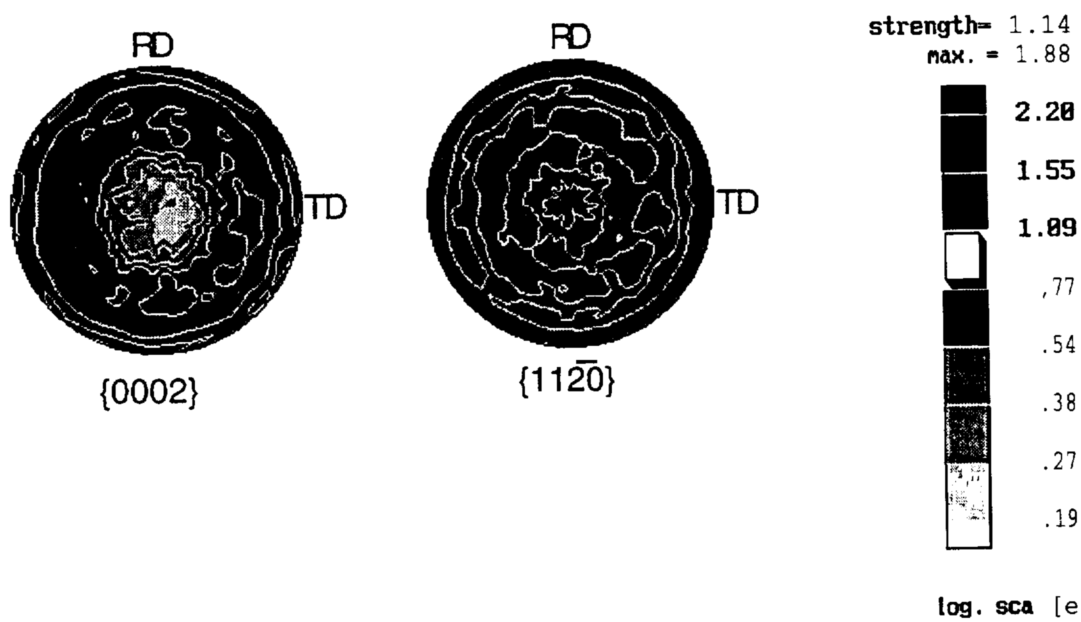


Figure 3. {0002} and {11 $\bar{2}$0} indirect pole figures for DC titanium strip. The casting direction is labelled RD on the pole figures.

The post CR textures for DC and IM foils are shown in Figure 4. After CR, the texture intensifies (-3 x random). In both foils, the c-axes were concentrated in the normal direction (ND) -transverse direction (TD) plane tilted approximately 35° from the ND towards the TD. Such split textures are commonly observed in IM titanium foils produced by cold rolling [7-10].

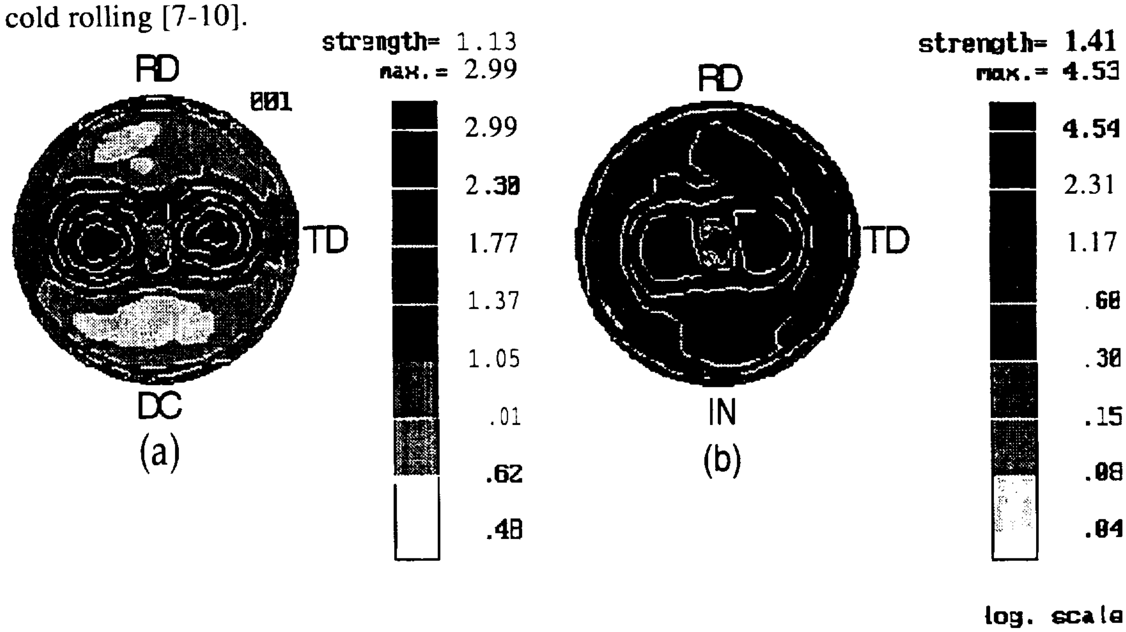


Figure 4. {0002} indirect pole figures for DC and IM foils.

Mechanical Properties

Mechanical properties results are presented in Table 2 along with the results obtained in a recent investigation of DC titanium alloys [2]. It is difficult to make an accurate comparison of mechanical properties due to the differences in composition between the DC and IM foils, however, some general observations can be noted. In agreement with Gaspar *et al.* [2], the room temperature yield stress (YS) and ultimate tensile stress (UTS) were relatively anisotropic in the IM titanium specimen. It was additionally observed that the IM specimens exhibited -40% greater tensile elongation's to failure than the DC specimens. Consistent with the observations of Gaspar *et al.* [2], YS and UTS were relatively anisotropic in the DC foils. The elongation, however, was noticeably lower in the transverse direction. They attributed this difference to crystalline anisotropy resulting from unidirectional rolling.

Summary and Conclusions

In general, this study has confirmed prior reports that high quality DC foils with mechanical properties comparable to IM foils can be successfully produced using the MORST process. Furthermore, these foils exhibit the same microstructure and textures as IM processed foils with lower interstitial impurity contents.

Table 2. Tensile Properties of CP-Ti Foils

Sample [reference]	Orientation	YS (MPa)	UTS (MPa)	ε_p (%)
IM [this study]	L	545±7	645±7	16.3±2.5
IM [this study]	T	515±35	620±28	16.8±1.1
IM CP-Ti [2]	L	339	459	31.5
IM CP-Ti [2]	T	362	438	28.7
DC [this study]	L	910	1010	151
DC [this study]	T	978±68	1170±17	7.3±1.4
DC Ti-1.25Al-0.8V [2]	L	762	841	18.6
DC Ti-1.25Al-0.8V [2]	T	807	887	7.4

L = longitudinal, T = transverse, IM = ingot metallurgy, DC = direct cast, base strain rate= 7.4 x 10^{-3} s^{-1}
NOTE: in reference [2], a nominal strain rate of 10^{-4} s^{-1} was used

Acknowledgments

The authors would like to thank Mr. W. D. Brewer at the NASA-Langley Research Center for supplying the DC titanium strips.

References

1. T. A. Gaspar and L. E. Hackman, *Materials Science and Engineering , 1991,* vol. A133, pp. 676-679.
2. T. A. Gaspar, T. A. Stuart, I. M. Sukonnik, S. L. Semiatin, E. Batawi, J. A. Peters and H. L. Fraser, Producing Foils From Direct Cast Titanium Alloy Strip, Ribbon Technology Corporation, Contractor Report #NASA CR-4742, May 1996 (1996).
3. T. A. Gaspar, L. E. Hackman, E. Batawi and J. A. Peters, *Materials Science and Engineering , 1994,* vol. A170/A180, pp. 645-648.
4. J. Kallend, U. F. Kocks, A. D. Rollett and H.-R. Wenk, *Materials Science and Engineering* , 1991, *vol.* A132, pp. 1-11.
5. M. V. Akdeniz and J. V. Wood, *Materials Science Forum , 1994,* vol. 157-162, pp. 1351-1356.
6. N. W. Blake and R. W. Smith, *Canadian Journal of Physics , 1982,* vol. *60,* pp. 1720-1724.
7. M. J. Philippe, F. Wagner and C. Esling, in Eighth International Conference on Textures of Metals, J. S. Kallend and G. Gottstein, Eds., The Metallurgical Society, Santa Fe, NM, 1988, pp. 837-842.
8. W. F. Hosford, in Oxford Engineering Science Series, A. L. Cullen, L. C. Woods, J. M. Brady, C. Brennen, W. R. E. Taylor, M. Y. Hussaini, T. V. Jones and J. V. Bladel, Eds., Oxford University Press, New York, 1993, vol. 32, pp. 132-134.
9. E. Tenckhoff, Deformation Mechanisms. Texture. and Anisotropy in Zirconium and Zircaloy, American Society for Testing and Materials, Philadelphia, PA, (1988), vol. STP 966.
10. D. R. Thornburg and H. R. Piehler, in Second International Conference on Titanium Science and Technology, R. I. Jaffee and H. M. Burte, Eds., Plenum Press, vol. 2, Cambridge, MA, 1973, pp. 1187-1197.

Page intentionally left blank

URC97158

A Novel Microcharacterization Technique in the Measurement of Strain and Orientation Gradient in Advanced Materials

NH 5 / 1

5157/39

H. Garmestani[1,2,3], K. Harris[3], L. Lourenco[2]

Center for Nonlinear & Nonequilibrium Aeroscience &
Florida A&M University
1800 E. Paul Dirac Rd., Tallahassee, FL 32306-4005

Department of Mechanical Engineering
FAMU-FSU College of Engineering
2525 Pottsdamer St., Tallahassee, FL 32310-2175

Center for Materials Research and Technology
Florida State University, Tallahassee, FL 32306-4000

Introduction

Representation of morphology and evolution of the microstructure during processing and their relation to properties requires proper experimental techniques. Residual strains, lattice distortion, and texture (micro-texture) at the interface and the matrix of a layered structure or a functionally gradient material and their variation are among parameters important in materials characterization but hard to measure with present experimental techniques, Currently techniques available to measure changes in interred material parameters (residual stress, micro-texture, microplasticity) produce results which are either qualitative or unreliable. This problem becomes even more complicated in the case of a temperature variation. These parameters affect many of the mechanical properties of advanced materials including stress-strain relation, ductility, creep, and fatigue. A review of some novel experimental techniques using recent advances in electron microscopy is presented here to measure internal stress, (micro)texture, interracial strength and (sub) grain formation and realignment. Two of these techniques are combined in the chamber of an Environmental Scanning Electron Microscope to measure strain and orientation gradients in advanced materials. These techniques which include Backscattered Kikuchi Diffractometry (BKD) and Microscopic Strain Field Analysis are used to characterize metallic and intermetallic matrix composites and superplastic materials. These techniques are compared with the more conventional x-ray diffraction and indentation techniques.

Experimental Techniques

EBSP and OIM Techniques

Electron backscattered pattern (EBSP) in the SEM originates from elastically backscattered and diffracted electrons. These electrons are formed when stationary primary electron beam is made to hit the surface of a specimen inclined at 70°, The diffraction patterns are imaged on a phosphor screen placed close to it, as illustrated in Figure 1. The phosphor screen is viewed through an optical port using a high gain television camera which in turn is interfaced to a computer. The resulting Kikuchi pattern (Figure 2) is recorded in the computer and indexed. By indexing successive patterns from hundreds of selected points on the sample surface, sufficient data can be collected to determine both macroscopic and

local orientation texture and to provide a detailed survey of nearest neighbor orientation relationships. Following upon the work of Venables[1], who developed the EBSP technique in the SEM, Dingley[2] advanced the technique, by interfacing with a computer to produce an on-line analysis of the diffraction patterns. This technique was later automated [3, 4] and evolved into the Orientation Imaging Microscopy (OIM) technique.

The OIM technique is essentially an extension of the electron backscattered patterns (EBSP) technique. Here, the EBSPs are collected from points on the sample surface over a regular grid and then automatically indexed.

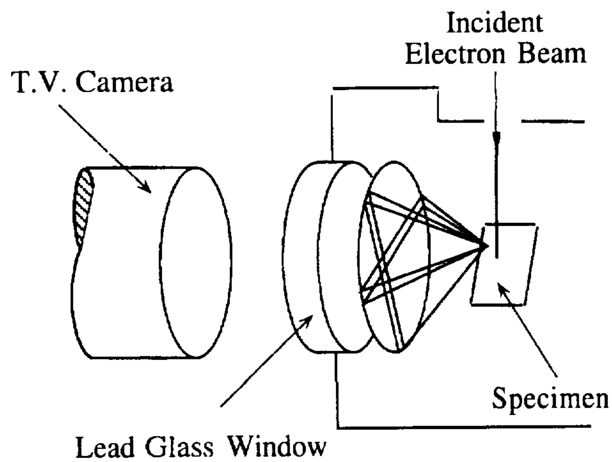


Figure 1- Schematics of BKD technique.

Figure 2- Typical Al-Kikuchi pattern

From this data, a map, called an OIM micrograph, is constructed displaying changes in crystal orientation over the specimen surface. In the OIM micrograph, the orientation of each point in the microstructure is known and hence the location, length and misorientation of all boundaries. This information is used to construct a micrograph based on a criteria input by the investigator. For example, a contiguous grain may be defined as a crystallographic entity on the basis that all points within it must have an orientation within 5 0 100, or 15°. Typical OIM micrographs require several hundred to several thousand EBSP measurements to be taken on a hexagonal grid of points, with spacing of 0.2 mm to several microns.

In addition to the orientation measured, an image quality (IQ) parameter which represents the sharpness of the electron backscatter patterns can be determined. This parameter is associated with the presence and intensity of local plasticity and other defects. An IQ micrograph can be produced by mapping the data from each pixel over the entire region of interest. The distribution of grain boundary misorientations (mesotexture) and pole figures can also be constructed from the data. This fully automated technique provides for the examination of the microtexture and mesotexture of large regions of the specimen.

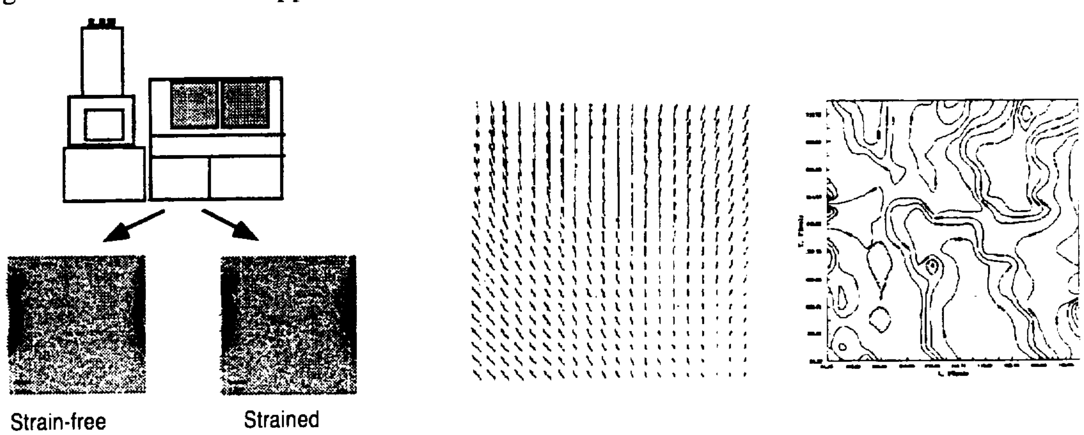
Microscopic Strain Field Analysis

We have developed a technique for microscopic strain field analysis (MSFA) which combines in situ heating or straining in an Environmental Scanning Electron Microscope (ESEM) with digital image processing [5]. This technique can be used to investigate the granular and intergranular strain localization during deformation. Figure 3 is a schematic which illustrates the application of MSFA to measure thermal strains in inhomogeneous materials. An ElectroScan model E-3 Environmental Scanning Electron Microscope

(ESEM) equipped with a heating and/or tensile stage, is used to image the specimen in a strain-free state and in a strained state. A digital micrograph of the region of interest is collected for each state. These images can be collected in real time (10 pictures per second) using a SGI computer interface. After the specimen is strained, the same feature must be centered before collecting the second micrograph.

The displacement field is determined by comparing the images of the strain-free and strained microstructure using the cross correlation approach. This approach has been discussed in detail for the measurement of particle displacements in the analysis of two dimensional velocity fields in fluid dynamics [6]. In brief, a small interrogation region where the material can be assumed to deform uniformly is selected. The displacement vector for the region is determined from the cross correlation of the equivalent regions in the two images. Two 2D Fourier transforms are applied to the interrogation region to obtain the two dimensional cross correlation function. In this two-dimensional array, the location of the maximum is proportional to the local displacement vector.

The cross correlation function is calculated for interrogation regions over the entire image to produce a map of. the local displacement. Displacements are displayed as an array of vectors. The displacement gradients and the subsequent components of strain are calculated using a finite difference approach.



1. ESEM in situ straining experiments, collection of digital images of strained and strain-free microstructures.

2. Cross correlation of digital micrographs to obtain local displacements.

3. Differentiation of displacements to obtain isostrain contours.

Figure 3- Schematic of MSFA for measurement of thermal strains in inhomogeneous materials.

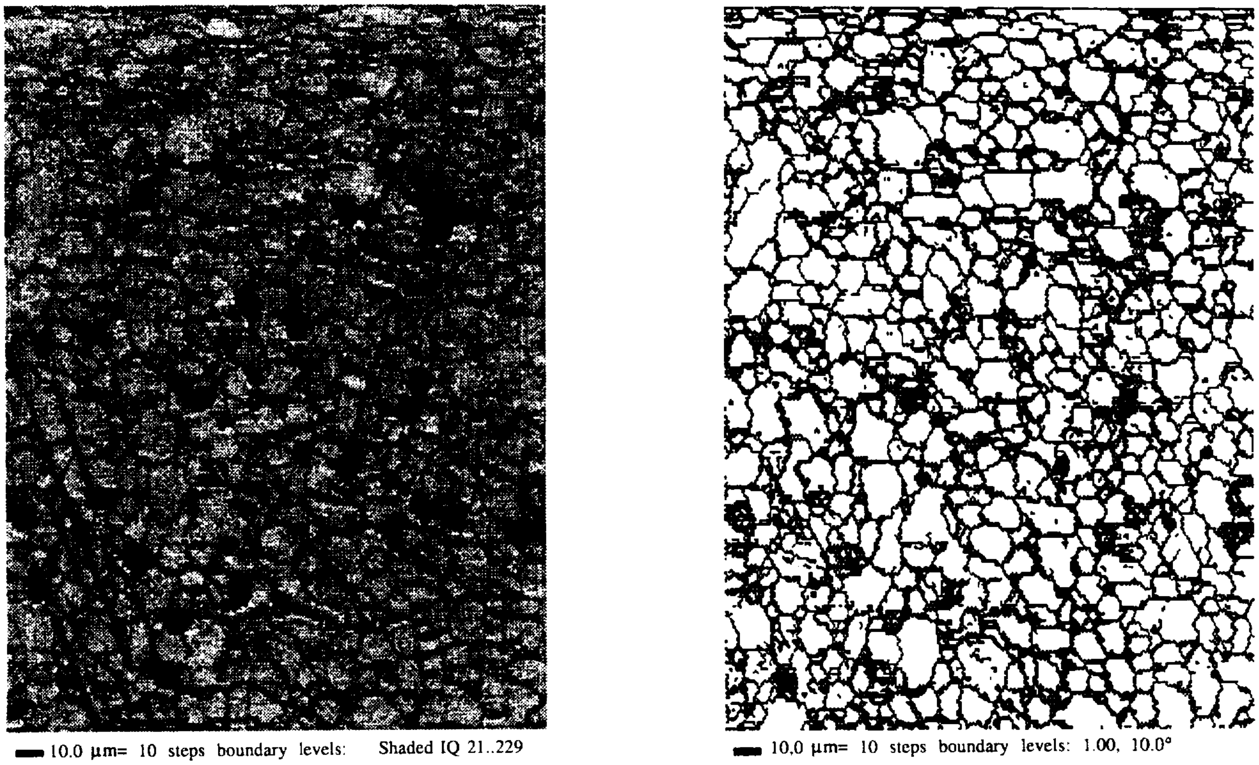
Results and Discussion

Materials and Materials development

In this section, the preliminary results of micromechanical characterization of a Al-8090 superplastic material is presented. Tensile specimens of the material were deformed at 516 “C and at a rate of 10^{-4} see-1 to different strain values and then characterized. The techniques described above were used for characterization, and this include Orientation Imaging Microscopy (OIM), Electron Backscattered Kikuchi Pattern (EBSP), load relaxation and strain rate change tests.

OIM Microstructural Analysis

Figure 4-a is the IQ micrograph for a sample deformed to 15%. The microstructure is essentially equiaxed and similar to that obtained by conventional optical or scanning electron microscopy techniques. Figure 4-b is an OIM image constructed using the same data set as



(a) (b)

Figure 4- a) Image Quality (IQ) micrograph and b) Disorientation boundary micrograph-Thin and thick lines represent grain boundaries with disorientation greater than 1° and greater than 10° respectively. Al-8090 specimen deformed to 15% strain.

Figure 5- Disorientation boundary micrographs. Black lines depict boundaries with misorientations greater than 3°.

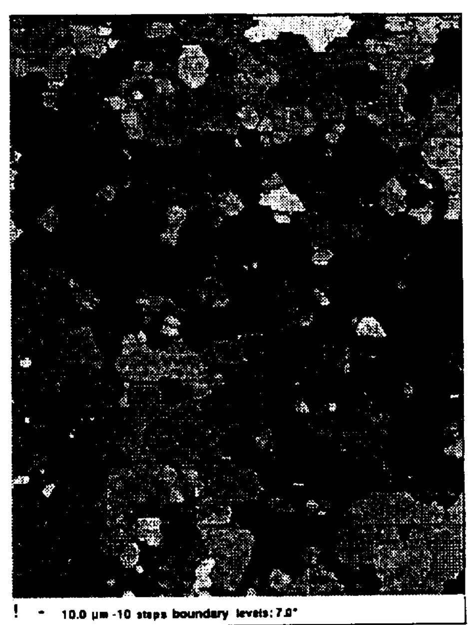
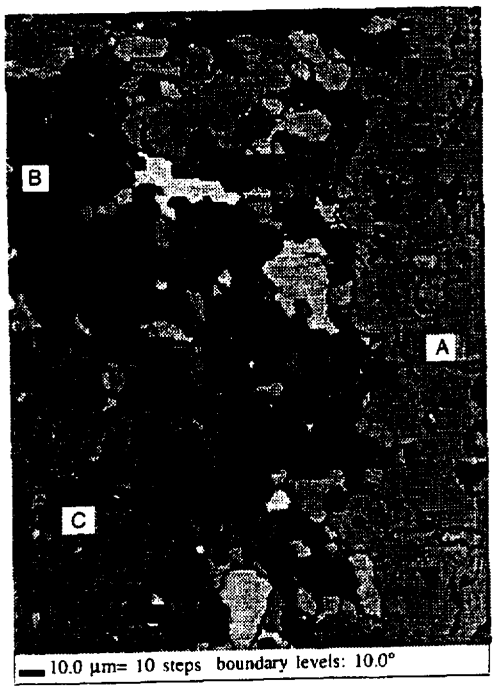
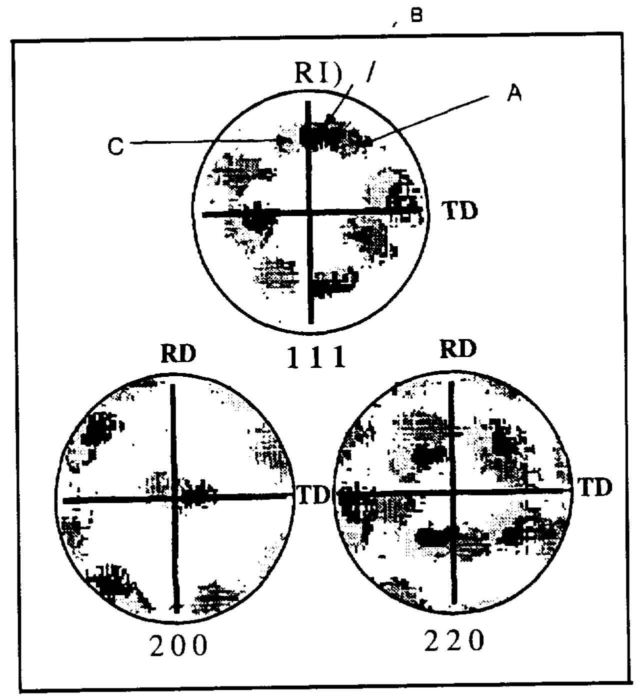


Figure 6- Misorientation boundary micro graphs. Black lines depict boundaries with misorientations greater than 7°.

in Figure 6-a but drawn to reveal boundaries with disorientation between 1° and 10° as thin lines, and those boundaries with misorientation greater than 10° as thick lines. The microstructure appears to be equiaxed. If boundary disorientation greater than 3° is

chosen, the micrograph of Figure 5 is produced. In this micrograph, regions of the same orientation within a tolerance of 3° are distinguished by a uniform color. Similar analysis is presented in Figures 6 and 7-a for boundary disorientation greater than 7°, and 10° respectively. When boundaries with misorientations greater than 10° are plotted, the microstructure consists of fine grains sandwiched between coarse grain structures (Figure 7-a). The corresponding EBSP pole figures obtained from these grains are shown in Figure 7-b. It is clear that the microstructure exhibits a rotated cube type texture with three distinct pole segments, A, B and C corresponding to the three grains A, B and C in Figure 7-a.





(a) (b)

Figure 7- (a) 100 Disorientation boundary micrographs. (b) Pole figure representation of the previous figures. Dark gray, black and light gray represent the three right, top right and bottom left coarse regions.

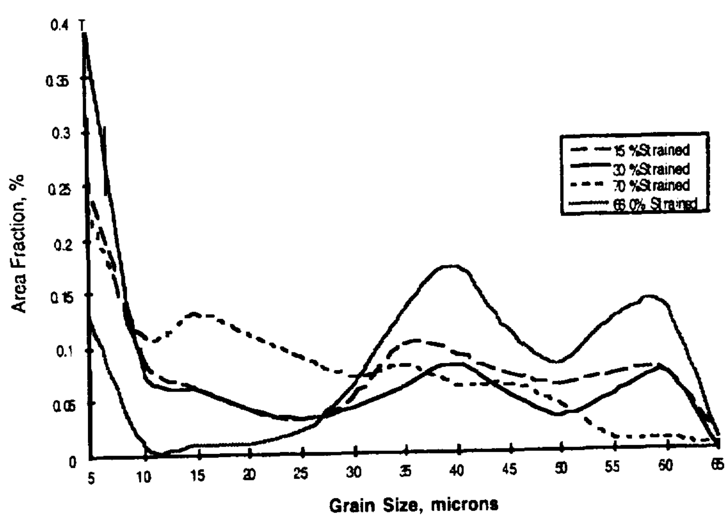


Figure 8- Composite of Test Results for Al-Li 8090 at (a) 516 ∞C and (b) Disorientation histogram for AlLi-8090 at Four Strain Levels.

It is evident that these techniques have opened up a new horizon in the art of microstructural characterization. It is well accepted that grain boundary sliding occurs at high angle grain boundaries. The mechanisms of such process are still under argument. The experiments performed above can be analyzed to provide information on the nature of the grain boundaries and distinguish the high angle grain boundaries which are candidates for grain boundary sliding. Figure 8 provides the result of grain boundary disorientation distribution for different ranges of strain rates,

Summary and Conclusions

A novel microcharacterization technique is introduced which measures strain and texture gradients and uses the chamber of an Environmental Scanning Electron Microscope. It combines and adopts the most important features of two other techniques of characterization, EBSP and PIV (Particle Image Velocimetry). In general, this novel experimental technique will open new opportunities in the characterization of materials. It is shown that it was possible to measure strain and strain gradients with a resolution of 0.2% and orientation within 0.3 degrees. At the present these techniques are applied to superplastic materials and composites materials. It is intended to extend the range of applicability of these techniques to Nanostructures of high strain rate superplastic material using a Transmission Electron Microscope (TEM).

Acknowledgments

The funding for this work was provided by NASA through Cennas and Center for for Materials Research and Technology (MARTECH) and also through a grant by the Department of Transportation.

REFERENCES CITED

[1] J. A. Venebles and C. J. Harland, in *Phil. Mag.*, vol. 27, 1973, pp. 1193-1200.

[2] K. Z. Baba-Kishi and D. J. Dingley, "Backscatter Kikuchi Diffraction in the SEM for Identification of Crystallographic Point Groups," Scanning, vol. 11, pp. 305-312, 1989.

[3] S. 1. Wright, B. L. Adams, and K. Kunze, "Application of a new Automatic Lattice Orientation Measurement Technique to Polycrystalline Aluminum," *Materials* Science *and Engineering,* vol. A160, pp. 229-240, 1993.

[4] H. Garmestani, P. N. Kalu, and D. Dingley, presented at Proceeding of the Light Weight Alloys for Aerospace Applications III, 124th Annual Meeting of TMS, 1995, Feb 12-16.

[5] K. Harris, H. Garmestani, and L. Lourenco, "Microscopic Strain Field Analysis for In Situ ESEM Investigation," *to be submitted to the Journal of Experimental Mechanics.*

[6] L. Lourenco and A. Krothapalli, "On the accuracy of velocity and vorticity measurements with PIV," *Experiments in Fluids,* vol. 18, pp. 421-428, 1995.

URC97159

5/68/37

Distributed Piezoelectric Element Method for Vibration Control of Smart Beams

Zhong L. Xu
(Fax : (210) 544-3802, E-mail: Zhong@utb1.utb.edu)

Abstract In this paper, a **new** method for active control beams — Distributed Piezoelectric Element (DPE) method is presented in which new method for designing piezoelectric modal sensor and modal actuator are given and the scheme for modal control is proposed. The observation spillover and control spillover related to this method are also analyzed and the approaches to eliminate or to reduce spillover are given.

The Distributed Piezoelectric Element Method has provided three advantages :

(1) the change of the order of controlled modes can be realized by software, but it does not need to change the shape of piezoelectric sensor and actuator laminae on the structures;

(2) it may not be necessary to change the sensor and actuator laminae for the change of the basic structures;

(3) the sensor/actuator could not only sense/actuate the local strain or strain rate, but also the entire strain or strain rate.

Therefore, this method is appreciated for the vibration control by using a on-board computer.

Key words smart beam; vibration control; distributed piezoelectric element; observation spillover; control spillover

I Introduction

In recent 10 years, smart structures for vibration control of the space flexible structures have attracted significant attention in the vibration control field. Applying smart structure to vibration control can be trace back to Beiley and Hubbard [2], who used PVDF as a distributed sensor and actuator in a cantilever beam to control its vibration. The following 10 years, there were many significcamt works [3-14].

Modal control methods are often used in vibration control which need distributed sensor/actuator to sense/actuate desired modes exactly. In this method, however, the class of full distributed sensor/actuator is not fit the change of structural shape and load, the change of the number of controlled modes and the change of the local deformation of structure . These evoke us to seek a method that may overcome the above shortcomings.

As we know, a single-piece piezoelectric sensor/actuator layer can only sense/actuate an average strain of the area it covers, and from which the local strain is hard to be determined, moreover, it may lead to observability/controllability deficiencies. Seeing that case, we address the idea, so-called " Entirely distribution, and separated treatment". That means, the entire piezoelectric sensor/actuator layers are cut into several independent pieces called sensor/actuator elements, and each of which can senses/actuates a local strain state. And also we may have information in a whole structure through synthetic treatment for the local information. Also, we address the distributed piezoelectric element (DPE) method for vibration control of beams, in which the observed modal coordinates and modal velocities in modal sensor are drawn from the output charge or current of sensor elements, while, the modal actuators are obtained through modulating the spatial distribution of voltage exerted on the actuator elements. Then use the modal sensor and modal actuator to realize the modal control of beams. Also, the control and observation spillover are analyzed and the improved methods are given.

II Differential equation of motion for a smart beam

Consider a slender elastic beam, two piezoelectric laminae, for instance PVDF, are bonded onto the upper face of the beam as actuator and the lower face of it as sensor respectively. (see Fig 1) In particular, for PVDF with single axis, when its principle axis coincides with the axis of the beam, the constitutive equations are

$$\sigma = Ys - e_{31}E_3 \qquad (1)$$
$$D_3 = e_{31}s + \varepsilon_{33}E_3 \qquad (2)$$

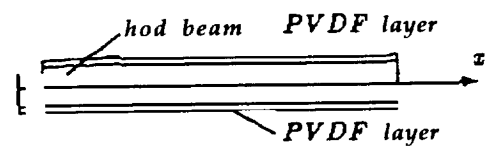


Fig.1 The smart beam with Piezoelectric layers

where σ, s, D_3, and E_3 are the stress, strain, electric displacement and electric field intensity, respectively Y, e_{31} and ε_{33} are the elastic constant, piezoelectric constant and dielectric constant, respectively

Based on the theory of Euler - Bernoulli beam, the differential equation of motion of the beam can be derived

$$\rho A(x)\frac{\partial^2 w}{\partial t^2} + \frac{\partial}{\partial x^2}\left(YJ(x)\frac{\partial^2 w}{\partial x^2}\right) = -b_3 e_{31}^{s}\frac{\partial^2 (r_s V)}{\partial x^2} \qquad (3)$$

where PA(x) is the equivalent mass density in a unit length of the beam. Eq. (3) represents that the motion equation of smart beam responds to voltage applied to actuator laminae, it is also called actuator equation.

III Modal sensor design

The electric field intensity exerted on sensor laminae is zero, the electric displacement D_3 can be obtained from Eq.(2). Integrating over the length of the beam yields the so-called sensor equation of the beam

$$q = -e_{31}^{s} b_1 \int_0^{l} \frac{1}{2}(z_0 + z_1)\frac{\partial^2 w}{\partial x^2}dx = -e_{31}^{s} b_1 \int_0^{l} r_s \frac{\partial^2 w}{\partial x^2}dx \tag{4}$$

where q is the charge of full sensor lamina. Differentiating Eq. (4) with respect to t, we have

$$I(t) = \frac{dq}{dt} = -e_{31}^{s} b_1 \int r_s \frac{\partial^2 w}{\partial x^2 \partial t}dx \tag{5}$$

which gives the value current of the beam's surface. Those are not able to obtain the information of modes. In order to draw the modal coordinates and modal velocities from the voltage output of the sensor lamina, the sensor lamina boned on the surface of the beam is separated into m_s sensor elements say $S_1, S_2, \cdots, S_{m_s}$, Δx_i is the length of the ith sensor element and x_i is the coordinate of center of the ith element, as shown in Fig, 2. From Eq. (4), the charge of the ith sensor element caused by strain of the beam is that

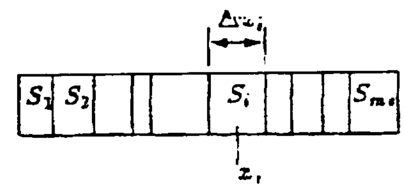


Fig.2 Segmentation of sensor layer

$$q_i = -e_{31}^{s} b_1 \int_{s_i} r_s \frac{\partial^2 w(x,t)}{\partial x^2} = b_1 e_{31}^{s} \int_{s_i} \varepsilon(x,t)dx\,, \quad i = 1, 2, 3, \cdots, m_s \tag{6}$$

Making mode truncation, the transverse displacement of the beam can be expressed as a linear superposition of the former m_s modes of the beam

$$w(x,t) = \sum_{j=1}^{m_s} \eta_j^{*}(t) W_j(x) \tag{7}$$

where $W_j(x)$ is the jth mode shape of the beam, $\eta_j^{*}(t)$ is the observed value of the jth mode coordinate. Then we have

$$q_i(t) = b_1 e_{31}^{s} \sum_{j=1}^{m_s} \eta_j^{*}(t) \int_{s_i} -r_s \frac{\partial^2 W_j(x)}{dx^2}dx = b_1 e_{31}^{s} \sum_{j=1}^{m_s} \eta_j^{*}(t) \int_{s_i} \varepsilon_j^{*}(x)dx\,, \quad i = 1, 2, \cdots, m_s \tag{8}$$

is the jth strain mode of sensor lamina. Eq. (8) can be rewritten in the matrix form

$$\{q\} = e_{31}^{s} b_1 [\varepsilon^{s}]\{\eta^{*}\} \tag{9}$$

where &s] is $m_s \times m_s$ strain mode matrix. The output charge of several neighboring elements can be superposed together. On the other hand, removing the output charge of a element is equivalent to removing this element. Therefore, the number, size and location of sensor elements can be changed through superposing and removing the output charge of sensor elements. Without loss of generality, the matrix [ES] is assumed to be non-singular, the m_s mode coordinates can be solved from Eq. (9):

$$\{\eta^{*}(t)\} = \frac{1}{b_1 e_{31}^{s}}[\varepsilon^{s}]^{-1}\{q(t)\} \tag{10}$$

If the sensor lamina is divided into many very small elements with equal length Ax, all the former m_s mode coordinates and mode velocities can be determined. In this case, Eq. (8) becomes

$$\{q(t)\} = e_{31}^{s} b_1 \Delta x [\varepsilon^{s}]\{\eta^{*}\} \tag{11}$$

Consider the orthogonally of strain modes and $[\varepsilon^{s}]$ is non-singular. From Eq. (11), the equation corresponding to Eq. (10) can be obtained

$$\{\eta^{*}(t)\} = \frac{1}{b_1 e_{31}^{s} \Delta x}[\varepsilon^{s}]^{-1}\{q(t)\} \tag{12}$$

so that the observed values, $\{\eta^{*}(t)\}$ and $\{\dot{\eta}^{*}(t)\}$, of former m_s mode coordinates and mode velocities can be obtained. The equations (10) or (11) and (12) show the process of mode observation performed by mode sensors, the observed values of former m_s mode coordinates and mode velocities can be obtained from output charge or current (differentiating q(t) and $\eta^{*}(t)$)of all sensor elements.

IV Modal actuator design

The piezoelectric mode actuator is designed by adjusting the distribution, in space, of voltage of actuator lamina. Multiplying both sides of Eq. (6) by W_j (X), integrating over the length of the beam with orthogonality of modes, we have

$$\ddot{\eta}_j(t) + \omega_j^2\eta_j(t) \cdot -b_3e_{31}^a \int_0 \frac{\partial^2(r_a V)}{\partial x^2} W_j(x)dx, \qquad j = 1, 2, \tag{13}$$

It can be seen that only some specified modes can be actuated by the voltage applied to piezoelectric actuator lamina through modulating the distribution, in space of the voltage. To this end, we design the voltage to be determined by

$$V(x,t) = \frac{1}{r_a}\sum_{j=1}^{n} p_j(t)\left(-YJ(x)\frac{d^2W_j(x)}{dx^2}\right) = \frac{1}{r_a}\sum_{j=1}^{n} p_j(t)M_j(x) \tag{14}$$

where n is the number of the order of required controlled modes, $n \le m_a$, To realize the distribution, in space, of voltage in Eq. (14), the piezoelectric actuator lamina is divided into ma elements, $A_1, A_2, \cdots, A_{m_a}$, $m_a > n$, the length and the coordinate of central point of the ith element is Δx_i and x_i respectively. Apply constant voltage to each element, the continuous distribution, in space, of voltage V(x, t) is approximated by "piecewise constant voltage". The applied voltage to the ith actuator element is

$$vi(t) = \frac{1}{\Delta x_i}\int_{A_i} V(x,t)dx = \frac{1}{r_{ai}\Delta x_i}\sum_{j=1}^{n} p_j(t)\int_{A_i} M_j(x)dx \tag{15}$$

where r_{ai} is the z-coordinate of the mid-plane of the ith actuator element and its matrix form is

$$\{V(t)\} = [C][\overline{M}]\{P(t) \tag{16}$$

P(t) represents mode control force vector. If the actuator elements are very small ones with same length Ax,

Eq.(16) becomes $$\{V(t)\} = \Delta X[C][\overline{M}]\{P(t)\} \tag{17}$$

Employing the orthogonality of modes yields

$$\begin{aligned} &\ddot{\eta}_k(t) + \omega_k^2\eta_k(t) = b_3e_{31}^a\omega_k^2 p_k(t), && k = 1, 2, \cdots, n \\ &\ddot{\eta}_k(t) + \omega_k^2\eta_k(t)\ 0, && k = n+1, \cdots \end{aligned} \tag{18}$$

that shows that the former n modes can be controlled. In this way, required voltage on each piezoelectric actuator element can be determined and the mode control can be approximately realized by applying these voltage to actuator elements.

V Modal Control

The modal control of the beam can be realized by designing the modal control forces $P(t)$ in modal equation of motion (18). Employing negative feedback for modal coordinates and velocities, the $P(t)$ are

designed as $$\{P(t)\} = -[G]\{\eta^*(t)\} - [H]\{\dot{\eta}^*(t)\} \tag{19}$$

where [G] and [H] are diagonal mode control gain matriced, $\eta^*(t)$ and $\dot{\eta}^*(t)$ are the former n modal coordinates and modal velocities observed from Eq. (10). Inserting Eq. (19) into Eq. (18) gives the closed loop equations

$$\ddot{\eta}_k(t) + \omega_k^2\eta_k(t) \cdot -b_3e_{31}^a\omega_k^2 g_k\eta_k^*(t)\ b_3e_{31}^a\omega_k^2 h_k\dot{\eta}_k^*(t) \qquad k = 1, 2, \ldots, n$$

Substituting Eq. (19) into Eq. (16) yields the voltage distribution

$$\{V(t)\} = -[C][\overline{M}[G]\{\eta^*(t)\} - [C][\overline{M}][H]\{\dot{\eta}^*(t)\} \tag{20}$$

Applying the ma voltages obtained from Eq. (20) on the actuator elements respectively can realizye the modal control of the beam.

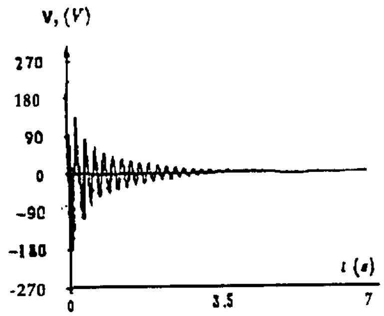


Fig. 6a Time behavior of the control voltage exerted on the 1 st actuator element

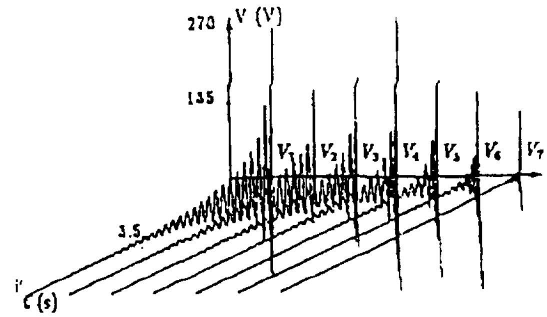


Fig. 6b Voltage distributions exerted on the 7 actuator elements

References

1 G L Anderson, Andrew Crawson and Jagdish Chandra. Introduction to Smart Structures. *Intelligent Structural Systems*, Edited by H S Tzou and G L Anderson, Kluwer Academic Publishers, 1992, pp1-8.

2 T Bailey, and J E Hubbard. Distributed Piezoelectric-Polymer Active Vibration Control of a Cantilever Beam. *Journal of Guidance and Controll*, Vol.8, No.5, 1985, pp650-611.

3 E F Crawley. Intelligent Structures for AerospaceA Technology Overview and Assessment. *AIAA Journal*, Vol.32, No.8, 1994, pp1689-1699.

4 E F Crawley end E H Anderson. Detailed Models of Piezoelectric Isotropic and Anisotropic Plates. *Journal of Intelligent Material Systems and Structures*, Vol.1, No.1, 1000, pp4-25.

5 Robert L Clark, Chris R Fuller, and A I Wicks. Characterization of Multiple Piezoelectric Actuators for Structural Excitation. *Journal of* Acoustical *Society of* America, Vol.90, No.1, 1991, pp346-357.

6 K Chandrashekhara and A N Agarwal. Active Vibration Control of Laminated Composite Plates Using Piezoeletric DevicesA Finite Element Approach. *Journal of Intelligent Material Systems And Structures*, Vol.4, 1993, pp496-508.

7 H S Tzou and M Gradre. Thoretical Analysis of a Multi-Layered Thin Shell Coupled With Piezoelectric Shell Actuators for Distributed Vibration Controls. *Journal of Sound and Vibration*, Vol.132, No.2, 1989, pp433-450.

8 H S Tzou. Active-Piezoelectric Shell Continua. *Intelligent Structural Systems*, Edited by H S Tzou and G L Anderson, Kluwer Academic Publishers, 1992, pp9-74.

9 G Shi and S N Atluri. Attire Control of Nonlinear Dynamics Response of Space-Frames Using Piezoelectric Actuators. *Computers and Structures*, Vol.34, 1990, pp549.

10 C C Won and J L Sulla. Application of Piezoelectric Devices to Vibration Suppression. *Journal of Guidance, Control and Dynamics*, Vol.17, No.6, 1994, pp 1333-1338.

11 C -K Lee. Piezoelectric Laminates: Theory and Experiment for Distributed Sensors and Actuators. *Intelligent Structural Systems*, Edited by H S Tzou and G L Anderson, Kluwer Academic Publishers, 1992, pp75-167.

12 C -R Lee and F C Moon. Modal Sensors and Actuators. *Journal of Applied Mechanics*, Vol.57, 1990, pp434-441.

13 C -K Lee, W -W Chiang and T C O'Sullivan. Piezoelectric Modal Sensor/Actuator Pairs of Critical Attire, Damping Vibration Control. *Journal of Acoustics Society of America*, Vol.85, No.7, 1991.

14 H S Tzou, J P Zhong and J J Hollkamp. Spatially Distributed Orthogonal Piezoelectric Shell Actuators: Theory and Applications. *Journal of Sound and Vibration*, Vol.177, No.3, 1994, pp363-378.

/// ... 5159/91

URC97160

Closed-Loop Aeromaneuvering for a Mars Precision Landing

Roy Smith
Electrical & Computer Engineering Dept.
University of California, Santa Barbara
Santa Barabara, CA 93106

Dhemetrios Boussalis and Fred Y. Hadaegh
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

Abstract

Controlled aeromaneuvering is considered as a means of achieving a precisely targeted landing on Mars. This paper presents a preliminary study of the control issues. The candidate vehicle is the existing Mars Pathfinder augmented with roll thrusters and a center of mass offset actuator. These allow control of both bank angle and lift force, giving the ability to control the range and cross-track during the aeromaneuvering entry. A preliminary control system structure is proposed and a design simulation illustrates significant targeting improvement under closed-loop control.

1 Introduction

This paper describes on-going work on the development of technologies needed for precision landings on Mars, The current Mars Pathfinder mission uses a small bluff vehicle on a ballistic trajectory. The error in specified landing position is an ellipse of approximately 300 km by 150 km.

The objective of this work is to develop the feedback control strategies necessary to reduce this error ellipse to less than 10 km, and, in the longer term, to less than 1 km. This level of precision will enable targeting for specialized science missions; for example: landing inside a crater of 20 km in diameter; or retrieving samples from a prior mission.

The major sources of landing error are navigation errors on the approach to atmospheric entry, and, more critically, uncertainties in the atmospheric properties. In order to reduce these errors it is necessary to use closed-loop feedback control during the atmospheric entry phase. The cent rolled aeromaneuvering issues are discussed in this paper. The complete mission will also involve parachute and terminal landing phases. Work on these phases is in the preliminary stages and is not discussed in detail here.

The candidate landing vehicle is based on the existing Pathfinder entry vehicle. This is a low lift vehicle, with maximum lift to drag ratio of less than 0.3, which imposes limitations on the control design problem. The advantages of using this vehicle are low development cost and access to an existing aerodynamic database. In its nominal configuration the Pathfinder vehicle generates no lift. We are proposing that a variable center of mass offset be used as an actuator to generate a controlled lift force and that roll thrusters be used to bank and thereby steer the vehicle.

1.1 Mission Scenario

This study considers the case where the atmospheric entry is made directly from interplanetary cruise, rather than first going into orbit. This is more appropriate for a small low cost mission, and gives a higher entry velocity. It does not consider the potential for reducing the navigational errors during an orbital phase.

The descent to landing consists of three phases: atmospheric aeromaneuvering; parachute; and terminal landing phase. The atmospheric entry phase begin at an altitude of 125 km when the vehicle has a velocity, V, of 7.5 km/second. The flight path angle, γ, is nominally -14.2 degrees. The parachute is deployed at an altitude of 8 to 10 km, when the velocity has decelerated to approximately 0.5 km/see. At the parachute deployment the fright path angle is between -45 and -60 degrees. The parachute is released at an altitude of approximately 5 km, when the velocity has further decelerated to 60 m/see. A three axis reaction control system (RCS) is then used to navigate and descend to a soft landing.

The work described here considers only the aeromaneuvering control aspect of this problem. The entry into the upper atmosphere is considered as the initial point for this control. The objective is to control the

atmospheric descent and guide the vehicle to the start of the parachute phase above a prespecified target on the planet.

The 1997 Pathfinder mission to Mars will use a bluff body on a ballistic entry trajectory. The errors in the atmospheric entry conditions (flight path angle, azimuth angle, and entry altitude) and uncertainties in the atmospheric properties and vehicle aerodynamics translate to a target accuracy of approximately 300 km by 100 km.

The objective of this work is to used closed-loop control during the atmospheric entry phase to reduce the error at parachute deployment to within 10 km or less. The controlled terminal phase will further reduce this error, with the longer term design objective being a total system landing error under 1 km.

1.2 Related work

Controlled aeromaneuvering has been considered in a number of other applications, Lifting trajectories have been used in the Apollo, Shuttle and Viking programs. See Dierlam [1], and the references therein, for more detail on the strategies used in these cases. Although the Viking program placed vehicles on Mars, these were not precision landings [2]. An open-loop strategy was used and the 3-σ ellipse was of the order of 120 km x 60 km. The Apollo program used a low L/D ratio (0.3) vehicle and bank angle control to maintain a reference drag profile. Crossrange requirements were met by changing the sign oft he bank angle in response to current crossrange error. The desired landing accuracy was 15 nautical miles.

The Shuttle program uses a vehicle with an approximate L/D ratio of 1.2 and has significantly greater accuracy requirements. Control is implemented via bank angle, angle-of-attack and a speed brake, which are used to fly reference drag and altitude rate profiles. In the case of Earth re-entry a better characterization of the atmosphere is available, reducing the uncertainty associated with the problem. In this case the entry conditions can also be more precisely specified.

Dierlam [1] describes a simulation study of a bank angle control strategy for landing a vehicle on Mars. A predictor-corrector strategy is used to predict the terminal errors which result from the currently estimated trajectory errors and forms the basis of the control algorithm. This work is similar in several respects although it considers a three degree of freedom simulation and does not consider the effects of uncertainty on vehicle orientation. The use of a predictor-corrector based guidance algorithm for aerobraking in the Martian atmosphere has been considered by Braun and Powell [3].

2 Vehicle Characteristics

The candidate vehicle is based on the Pathfinder aeroshell design, depicted in Figure 1. Detailed aerodynamic and configuration information is available in the work of Spencer and Braun [4], The vehicle studied here is augmented with two control actuators. The first is a variable center of mass offset, and is used to alter the trim angle-of-attack and thereby generate lift. The second is a pair of thrusters used to roll the vehicle, and thereby steer the lift force.

The prior work on the aerodynamic properties of the Pathfinder aeroshell [4] assumes a trim angle-of-attack, α, of zero which is the case for ballistic entry trajectories. We consider using a variable center of mass offset as an actuator, which necessitates developing an aerodynamic model for other values of α. This approach is based on a Newtonian impact theory model. This model is appropriate for the hypersonic entry considered here, and we also use it as an approximate ion for lower velocities closer to the planet surface. Future work will require the development of more detailed aerodynamic models for the lower velocities,

The lift and drag forces acting on the vehicle are given by, $L = QSC_L$ and $D = QSC_D$, where S is a reference area (in this case 5.515 m^2) and Q is the dynamic pressure, given by, $Q = \rho V^2/2$. The atmospheric density is ρ, and V is the vehicle velocity. The lift and drag coefficients, C_L and C_D, are function of the angle-of-at tack, α, and the side-slip angle, β. Because of the circular symmetry of the aeroshell we can combine α and β into a single effective angle of attack variable, a,, defined by

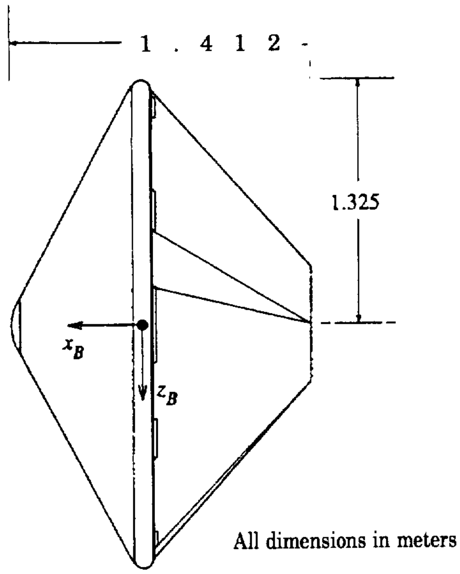
$$\sin^2\alpha_e = \cos^2\beta \sin^2\alpha + \sin^2\beta.$$

This is effectively the angle-of-attack in a frame rotated by an angle,

$$6' = \arctan\left(\frac{-\cos\alpha\sin\beta}{\sin\alpha}\right),$$

with respect to the wind frame. In this frame the lift and drag coefficients can be approximated by,

$$C_L = C_N\cos\alpha_e - C_A\sin\alpha_e \qquad \text{and} \qquad C_D = C_N\sin\alpha_e + C_A\cos\alpha_e,$$



Mass (kg)	552.0		
Inertias (kg.m^2)	I_x	I_y	I_z
	230.0	184.0	180.0
	I_{yz}	I_{xy}	I_{zx}
	0.0	0.0	0.0

Figure 1: Pathfinder entry vehicle configuration

where C_N and CA are the normal and axial coefficients, which are approximated by,

$$C_N = 0.4381\alpha_e \qquad \text{and} \qquad C_A = 1.7204 - 1.5623\alpha_e^2.$$

The pitching and yawing moments are given by, $M = QcSC_m$ and N = $QbSC_n$, where b and c are reference lengths, which are both 1.325 m in this case. These moment equations determine the static trim of the vehicle. In the standard configuration this is α = O, β = O. The actuation considered here involves offsetting the center of mass in the body Z axis, by an amount dz. With this offset the coefficients are,

$$C_m = C_{moz}dz + C_{m\alpha}\alpha + C_{mq}q \qquad \text{and} \qquad C_n = -C_{m\alpha}\beta + C_{nr}r.$$

where q is the pitch rate and r is the yaw rate. The coefficients are C_{moz} = 1.6757, $C_{m\alpha}$ = –0.5275, C_{mq} = –0.05, and C_{nr} = –0.054. Note that the Z axis center of mass offset does not affect the yaw moment. These Figure 2a) illustrates C_m and C_n as a function of α and β. This figure can be interpreted by considering the velocity vector as pointing directly at the viewer from the origin. The stem of each arrow is placed on the vehicle nose and the arrow then gives the relative size and direction of the corresponding moment acting on the vehicle. The figure shows dz = –O .099 which gives a static trim of α = –18.0 degrees (sin α = –0.309). This is the maximum magnitude trim angle and gives a maximum lift coefficient of C_L = 0.3531 and a lift/drag ratio of 0.2305.

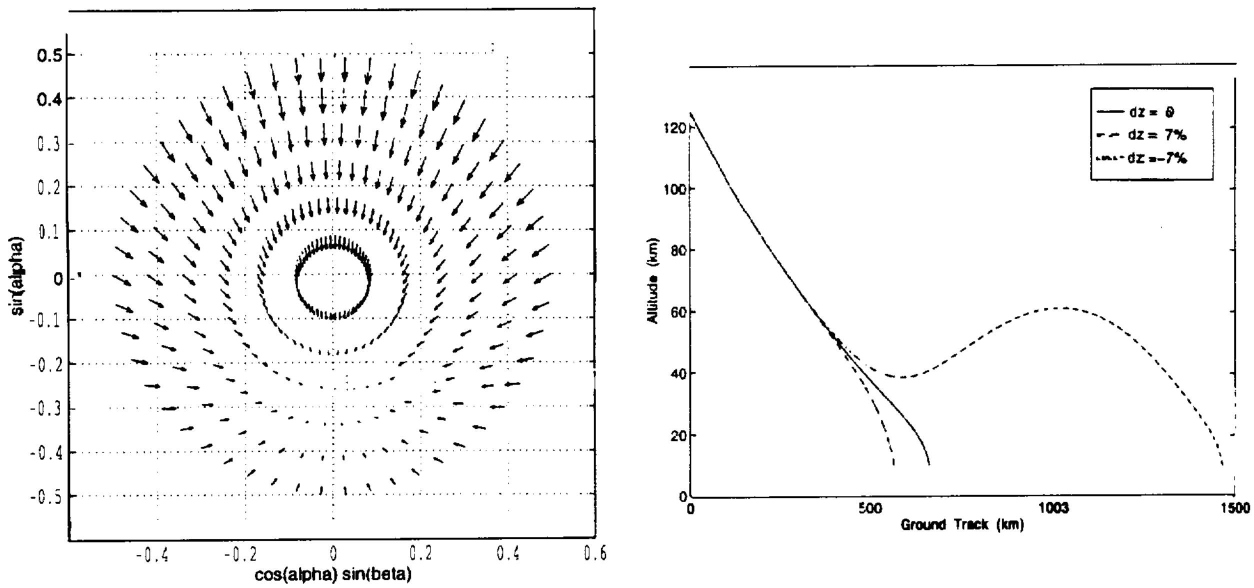
The achievable range is illustrated in Figure 2b). The entry trajectories are illustrated for the maximum and minimum angles-of-attack. This corresponds to a movement of the center of mass of ±7% of the vehicle radius. The resulting ground track is extended over 800 km beyond the ballistic case. Varying the center of mass offset, dz, allows control actuation over this range.

3 Control Issues

The two major sources of error are navigation errors at entry and atmospheric uncertainty. The most significant entry condition errors are in altitude (±2.5 km) and flight path angle (+1.0 degrees). The knowledge error associated with these is an order of magnitude less which gives the potential to correct for these errors during the aeromaneuvering phase. This work uses a simple exponential model for the Martian atmospheric density,

$$p = \rho_r e^{-0.1(R-R_r)}, \qquad R_r = 3,4290.0 \text{ km}, \qquad \rho_r = 7.8 \ x \ 10^6 \text{kg/km}^3,$$

as a function of altitude, R, expressed as a radius from the planet center. The uncertainty in atmospheric density considered in this work is ±25%.



a) b)

Figure 2: CM actuated vehicle: a) Pitch/ yaw moments. b) Trajectories for α = +18, O, -18 degrees

A more significant control issue is the nonlinearity in the control actuation. As illustrated previously, the lift force can be controlled by moving the center of mass. The actuator effectiveness is directly proportional to the dynamic pressure, Q. Figure 3 illustrates the lift and drag forces that result from a nominal trajectory of $\alpha = -13$ degrees.

At the atmospheric entry (t = 0), the density is so small that Q is effectively zero. As the density increases the V^2 term causes a rapid increase in Q. The vehicle decelerates quickly, causing Q to drop to a low level for the duration of the entry. There is a window of maximum effective control opportunity between t = 50 and t = 100 seconds. Only limited control capability is possible after t = 100 seconds. The vehicle position, in planet latitude and longitude, is very sensitive to both control actions and atmospheric perturbations occurring within that 50 second window.

To illustrate the nature of the system dynamics, we present a simplified set of dynamical equations for this problem. The velocity, V, *is* given by

$$\frac{dV}{dt} = \frac{-D}{m} - g\sin\gamma,$$

where g is the gravitational acceleration and m is the vehicle mass. The flight path angle, γ, equation is

$$V\frac{d\gamma}{dt} = \frac{L}{m}\cos\sigma - \left(g - \frac{V^2}{R}\right)\cos\gamma.$$

The azimuth/heading angle, ψ, is determined by,

$$V\frac{d\psi}{dt} = \frac{L}{\mathrm{m}\cos\gamma}\sin\sigma - \frac{V^2}{r}\cos\gamma\cos\psi\tan\mu,$$

where μ is the longitude. The cent rol system can influence the lift and drag (L and D above) by changing the center of mass offset, dz, and thereby changing α. The roll thrusters allow the controller to var y the bank angle, σ. Note that in the above equations all terms are actually functions of time. Not shown are the vehicle moment equations, or the equations determining vehicle position in planet centered coordinates.

4 Candidate Control Structure

Figure 4 illustrates the control structure considered for this problem. This consists of the following components.

Nominal Path. The nominal control actions, u_{nom}, are determined by an offline optimization. The issues that must be considered include: vehicle dynamics and maneuvering constraints; entry location

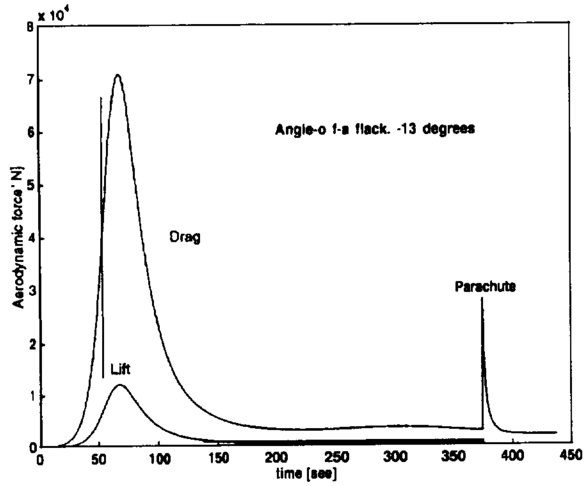


Figure 3: Lift and drag profiles

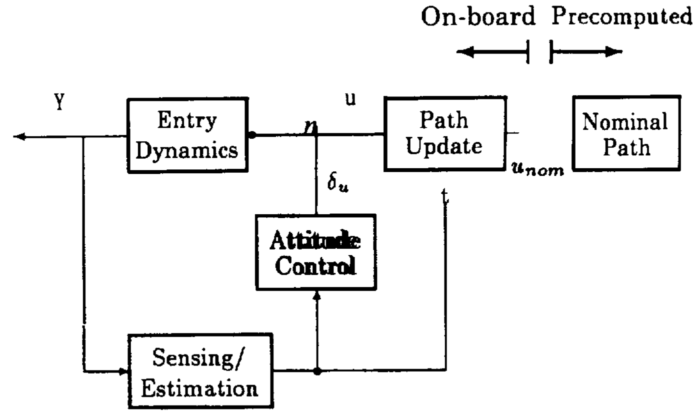


Figure 4: Control structure

and errors; control system achievable performance; atmospheric models; actuation saturation (rate and magnitude) constraints; terminal objectives (velocity, position, flight path angle); sensing limitations; robustness to system and atmospheric uncertainties; dynamic loading constraints; and thermal constraints.

Path Up date.' Trajectory deviations will be caused by entry condition errors and atmospheric perturbations and will require the calculation of a corrected trajectory. This involves a nonlinear prediction of terminal conditions based on the best current estimates of position, velocities, attitude and atmospheric variables, followed by a correction algorithm to determine the require control, u. Operational constraints (saturations, thermal limits, dynamic load limits) on the modified flight path must be considered. A sampled prediction/correction algorithm gives periodic up dates for u.

Attitude Control. Low level attitude control is required to fly the current modified flight path. The control input, u, is specified in terms of a (angle-of-attack) and σ (bank angle). A high bandwidth (continuous or digital) control law is required to generate the appropriate center of mass offset and roll thruster commands. A nonlinear/gain scheduled control is needed to account for widely varying vehicle response over the flight path.

Sensing/Estimation. The vehicle position, attitude and velocity is estimated from inertial measurement unit and any measurements (e.g. stagnation pressure). The parameters in atmospheric models, including density, are estimated and used for both attitude control and flight path updating. This is a highly nonlinear estimation problem, particularly for the atmospheric parameters. The quality of the velocity/attitude estimation is a function of IMU drift and will degrade over time. The quality of the atmospheric parameter estimates will improve over time.

5 Preliminary Design

A preliminary design has been evaluated and the results are shown graphically in Figure 5. The calculations involve a full 6 degree-of-freedom model. The primary reference for this type of model is Etkin [5], and for specific detail on the aeromaneuvering dynamic equations see Boussalis [6] and Smith [7].

The nominal trajectory was generated by selecting u_{nom} to be $\alpha = -13$ degrees and $\sigma = 0$ degrees. These values take the vehicle to close to the center of its achievable range (refer to Figure 2b).

The path update used a predictor to estimate the terminal range as a function of α. This predictor used a nominal model for the atmospheric density. The prediction result was used to select a new value for α. The desired bank angle was generated by a tracking controller following a precalculated azimuth trajectory. The estimator design problem is currently being investigate ed in more detail. The preliminar y design shown here used a measurement of the vehicle velocity, position, and attitude. This is optimistic in that these values will be degraded in an estimator. However, the use of an estimator would also give the opportunity to use more accurate atmospheric density values. The exact nature of this trade-off will be investigated in future research.

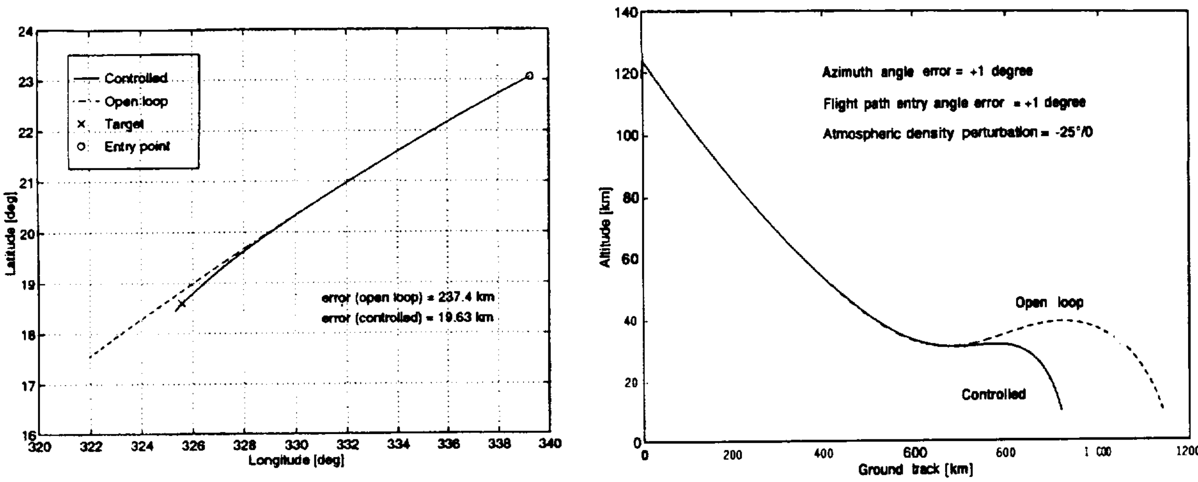


Figure 5: Simulated aeromaneuvering control system

The simulated entry conditions include a one degree error in both the azimuth and the flight path angle. In addition, the atmosphere was 25% less dense than the nominal case. Without closed-loop control, these conditions lead to a terminal landing error of 237 km. The closed-loop control reduces this error to under 20 km.

6 Summary

The simulation results indicate that controlled aeromaneuvering, using roll thrusters and center of mass offset actuation, has the potential to significantly reduce the landing error ellipse. On-going work is focusing on the development of higher performance controllers incorporating atmospheric density estimators. Overall system performance will be more extensively evaluated using more detailed models of system sensors, actuators, noise and drift. The attitude stabilization and RCS with optical sensing design problems will be studied for the parachute and terminal landing phases.

Acknowledgements

The research described in this paper has been performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

[1] T. A. Dierlam, "Entry Vehicle Performance Analysis and Atmospheric Guidance Algorithm for Precision Landing on Mars," 1990. MS Thesis, Massachusetts Institute of Technology.

[2] E. A. Euler, G. L. Adams, and F. W. Hopper, "Design and Reconstruction of the Viking Lander Descent Trajectories,)' *J. Guidance & Control, vol.* 1, September-October 1978.

[3] R. D. Braun and R. W. Powell, "Predictor-Corrector Guidance Algorithm for Use in High-Energy Aero-Braking Studies," *AIAA J. Guidance, Control* and Dynamics, vol. 15, pp. 672-678, May-June 1992.

[4] D. A. Spencer and R. D. Braun, "Mars Pathfinder Atmospheric Entry Trajectory Design," in *AAS/AIAA Astrodynamics Specialist Conf.*, August 1995. AAS Paper # 95-379.

[5] B. Etkin, *Dynamics of Atmospheric Flight,* Wiley & Sons, 1972.

[6] , "Investigation of the Longitudinal Motion of Low-Lift Entry Vehicles," Tech. Rep. EM 3456-96-002, Jet Propulsion Laboratory, May 1996.

[7] R. Smith, "Closed-Loop Aeromaneuvering for a Mars Precision Landing," Tech. Rep. Center for Control Engineering & Computation, University of California, Santa Barbara, Feb. 1996. CCEC-96-0205.

URC97161 NASA -1666/18

Formation Flying Control of Multiple Spacecraft

F. Y. Hadaegh and Kenneth Lau
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

P. K. C. Wang
Department of Electrical Engineering
University of California
Los Angeles, CA 90024

Abstract

The problem of coordination and control of multiple spacecraft (MS) moving in formation is considered. Here, each MS is modeled by a rigid body with fixed center of mass. First, various schemes for generating the desired formation patters are discussed, Then, explicit control laws for formation-keeping and relative attitude alignment based on nearest neighbor-tracking are derived. The necessary data which must be communicated between the MS to achieve effective control are examined. The time-domain behavior of the feedback-controlled MS formation for typical low-Earth orbits is studied both analytically and via computer simulation. The paper concludes with a discussion of the implementation of the derived control laws, and the integration of the MS formation coordination and control system with a proposed inter-spacecraft communication/computing network.

1 Introduction

Future use of multiple micro-spacecraft moving in formation for space exploration, constellation space antennas and interferometers, and space-based global communication systems calls for novel approaches to the design of spacecraft control systems. In particular in recent years, growing emphasis is placed on the concept of separated spacecraft interferometry (SSI). The SSI concept envisioned the collecting apertures to be located on separate spacecraft while central combining instruments to be located on yet another spacecraft. A virtual structure is therefore developed without the real need for maintaining the necessary structural rigidity. The SSI provides measurements unachievable with other techniques and allows long baseline lengths and orient ation changes.

This paper focuses on the development of a control system architecture for the coordination and cent rol of a fleet of micro-spacecraft moving in formation. Here, we are dealing with a collection of systems which interact with each other in a cooperative manner to achieve a common objective. Although control of a single spacecraft is based on well-established control theory concepts and methodologies, the control systems for multiple spacecraft moving in formation require architectures which differ from those of conventional single spacecraft control systems. To provide a desired formation, basic mathematical models for controlled movement of rigid bodies in free space is presented. The control laws for the coordination of spacecraft attitude during motion to achieve a specified objective (e.g. orient each spacecraft along a given direction) is also developed. This is followed by the derivation of control laws for formation keeping and relative attitude alignment, The time-domain behavior of the feedback-controlled formation flying for typical low-Earth orbits is studied both analytically and via computer simulation. Emphasis is placed on determining the information exchange needed for achieving and maintaining a desired formation, and the conditions

for ensuring formation stability in the presence of various types of perturbations. Other important factors such as collision-avoidance, spacecraft failures, loss of communication between spacecraft, and also various constraints imposed by implementation and the physical size of the micro-spacecraft are discussed.

2 Modeling of Multiple Spacecraft in Formation

To simplify the development, we consider only a spacecraft triad modeled by rigid bodies with fixed centers of mass moving in free space under the influence of a gravitational field and external disturbances. We introduce the following coordinate systems in the three-dimensional Euclidean space $\Re^3$: (i) an inertial coordinate system $\mathcal{F}_0$ with orthonormal basis $\mathcal{B}_0 = \{\mathbf{e_X}, \mathbf{e_Y}, \mathbf{e_Z}\}$, and (ii) a set of moving coordinate systems $\mathcal{F}_i, i = 1,2,3$ whose origins O_i are at the mass centers of the spacecraft. Let $\mathcal{B}_i = \{e_{ix}, e_{iy}, e_{iz}\}$ denote an orthonormal basis associated with the moving coordinate system (abbreviated by "MCS" hereafter) $\mathcal{F}_i$, and $[\mathrm{w}]_i$ the representation of the vector w with respect to basis $\mathcal{B}_i$. The the basis vectors in 13_0 and $\mathcal{B}_i$ are related by a linear t ransformat ion $\mathrm{C_i}$ defined by

$$e_{ix} = C_i e_x, \quad e_{iy} = C_i e_y, \quad e_{iz} = C_i e_z, \quad i = 1,2,3 \tag{1}$$

whose representation with respect to basis $\mathcal{B}_0$ is given by the direction cosine matrix $C(q_i) = (q_{i4}^2 - \hat{q}_i^T \hat{q}_i)I + 2\hat{q}_i\hat{q}_i^T - 2q_{i4}Q(\hat{q}_i)$, where

$$Q(\hat{q}_i) \triangleq \begin{bmatrix} 0 & -q_{i3} & q_{i2} \\ q_{i3} & 0 & -q_{i1} \\ -q_{i2} & q_{i1} & 0 \end{bmatrix}; \quad \hat{q}_i \triangleq \begin{bmatrix} q_{i1} \\ q_{i2} \\ q_{i3} \end{bmatrix}, \tag{2}$$

and $q_i = [\hat{q}_i^T, q_{i4}]^T$ denotes the unit quaternion with q_{ij} being the Euler symmetric parameters [5] defined by $q_{ij} = \varepsilon_{ij} sin(\phi_i/2), j = 1,2,3;$ $q_{i4} = cos(\phi_i/2)$, satisfying the constraint $\Sigma_{j=1}^4 q_{ij}^2 = 1, i = 1,2,3$ where ϕ_i is the principal angle and the ε_{ij}'s are the components of the principal vector of rotation ℓ_i defined by $\ell_i = \varepsilon_{i1} e_x + \varepsilon_{i2} e_y + \varepsilon_{i3} e_z = \varepsilon_{i1} e_{ix} + \varepsilon_{i2} e_{iy} + \varepsilon_{i3} e_{iz}$. The time derivitive of q_{ij} is related to the angular velocity $\omega_i = \omega_{ix} e_{ix} + \omega_{iy} e_{iy} + \omega_{iz} e_{iz}$ of $\mathcal{F}_i$ relative to the inertial coordinate system $\mathcal{F}_0$ by

$$\left.\begin{array}{rcl} \frac{d\hat{q}_i}{dt} & = & (q_{i4}\omega_i \quad \omega_i \times \hat{q}_i)/2, \\ \frac{dq_{i4}}{dt} & = & -(\omega_i, \hat{q}_i)/2, \end{array}\right\} \tag{3}$$

where w x v and w $\cdot$ v denote respectively the cross and scalar products of vectors w and v in $\mathcal{R}^3$.

Let d/dt_o and d/dt_i denote the time derivative operators with respect to $\mathcal{F}_0$ and $\mathcal{F}_i$ respectively; and $\mathcal{D}_i$ the time derivative operator d/dt_o in $\mathcal{F}_i$ defined by $\mathcal{D}_i \mathrm{w} \triangleq d\mathrm{w}/dt_i + \omega_i x w$ for w ε 73^3. The angular velocities of the i-th spacecraft or $\mathcal{F}_i$ relative to $\mathcal{F}_0$ are given by the following Euler's equations relating that time derivative of the angular momentum $I_i\omega_i$ with respect to $\mathcal{F}_0$ to the control torque τ_{ci}:

$$\mathcal{D}_i(I_i\omega_i) \triangleq d(I_i\omega_i)/dt_i + \omega_i \times (I_i\omega_i) = I_i d\omega_i/dt_i + \omega_i \times (I_i\omega_i) = \tau_{ci} \tag{4}$$

for $i = 1,2,3$, where I_i is the tensor of inertia associated with the i-th spacecraft. The time derivative of ω_i may also be taken with respect to $\mathcal{F}_o$.

In formation acquisition and keeping, we are interested in the relative motion between any pair of spacecraft labelled by subscripts i and j. Let the MCS $\mathcal{F}'$ be $\mathcal{F}_i$ as defined earlier (See Fig. 1), and

ρ_{ji} denote the position vector of the j-th spacecraft relative to $\mathcal{F}_i$. The evolution of ρ_{ji} with time is governed by

$$\mathcal{D}_i^2(\boldsymbol{\rho}_{ji}) = (\boldsymbol{f}_{cj} + \boldsymbol{f}_{gj})/M_j - (\boldsymbol{f}_{ci} + \boldsymbol{f}_{gi})/\boldsymbol{M}_i, \tag{5}$$

where $\mathcal{D}_i^2$ denotes the time derivative operator d^2/dt_o^2 in the MCS $\mathcal{F}_i$ given by

$$\mathcal{D}_i^2 \triangleq d^2/dt_i^2 + (d\boldsymbol{\omega}_i/dt_i) \times \cdot + 2\boldsymbol{\omega}_i \times d/dt_i + \boldsymbol{\omega}_i \times (\boldsymbol{\omega}_i \times \cdot). \tag{6}$$

3 Formation-Keeping Control Laws

The movement of spacecraft triad in formation can be achieved in marry ways. Here, we use a simple approach proposed in [6]. Let the first spacecraft be designated as the *leader*, and the remaining spacecraft as *followers.* The leader provides the reference motion $\boldsymbol{r}_1 = \boldsymbol{r}_1(t)$ (relative to $\mathcal{F}_o$) for the followers, and also two nonzero deviation vectors hi(t), $i = 2,3$, such that the point set $\mathcal{P}(t) = \{\boldsymbol{r}_1(\text{t}), \boldsymbol{r}_1(t) + \boldsymbol{h}_2(t), \boldsymbol{r}_1(t) + \boldsymbol{h}_3(t)\}$ defines the *desired formation pattern* at time t. It is assumed that the convex hull of P $(t)$is a simplex. We define the positional error for the i-th follower relative to the leader as $\boldsymbol{E}_i(t) \triangleq \boldsymbol{h}_i(t) - \rho_{i1}(t), i = 2,3.$

In what follows, we shall derive formation-keeping control laws assuming that each spacecraft in the triad moves in a low Earth orbit (LEO), and the desired motions for the followers are given by $\boldsymbol{d}_i(t) = \boldsymbol{r}_i(t) + \boldsymbol{hi(t)}$, $i = 2,3$. Using (5), we obtain the following differential equation for $\boldsymbol{E}_i$:

$$\begin{aligned} \mathcal{D}_i^2(\boldsymbol{E}_i) &\triangleq \ddot{\boldsymbol{E}}_i + \dot{\boldsymbol{\omega}}_i \times \boldsymbol{E}_i + 2\boldsymbol{\omega}_i \times \dot{\boldsymbol{E}}_i + \boldsymbol{\omega}_i \times (\boldsymbol{\omega}_i \times \boldsymbol{E}_i) \\ &= \mathcal{D}_i^2(\boldsymbol{h}_i) + \boldsymbol{f}_{g1}/\boldsymbol{M}_1 - \boldsymbol{f}_{gi}/\boldsymbol{M}_i + \boldsymbol{u}_{c1} - \boldsymbol{u}_{ci}, \end{aligned} \tag{7}$$

where $\boldsymbol{u}_{cj} = \boldsymbol{f}_{cj}/\boldsymbol{M}_j$.

Assuming a central Newtonian gravitational force field, the gravitational force acting on the i-th follower has the form: $\boldsymbol{f}_{gi} = -\mu \boldsymbol{M}_i \boldsymbol{r}_i/\|\boldsymbol{r}_i\|^3$, where μ is the geocentric gravitational constant; $\boldsymbol{r}_i$ is the vector specifying the position of the mass center of the i-th spacecraft relative to the inertial frame $\mathcal{F}_o$; and $\|\boldsymbol{r}_i\|$ the Euclidean norm of $\boldsymbol{r}_i$. Let $\boldsymbol{e}_{ri} = \boldsymbol{r}_i/\|\boldsymbol{r}_i\|$. We can write

$$\begin{aligned} \boldsymbol{r}_i &= \|\boldsymbol{r}_i\|\boldsymbol{e}_{ri} \\ &= \|\boldsymbol{r}_i\|\{(\boldsymbol{e}_{ri} \cdot \boldsymbol{e}_{ix})\boldsymbol{e}_{ix} + (\boldsymbol{e}_{ri} \,.\, \boldsymbol{e}_{iy})\boldsymbol{e}_{iy} + (\boldsymbol{e}_{ri} \,.\, \boldsymbol{e}_{iz})\boldsymbol{e}_{iz}\}, \end{aligned} \tag{8}$$

where $\boldsymbol{r}_j = \boldsymbol{r}_i + \boldsymbol{\rho}_{ji}$. For LEO spacecraft triad moving in formation at approximately the same altitude, we have $\|\boldsymbol{r}_i\| / \|\boldsymbol{r}_1\| \cong 1$. Thus, the components of $(\boldsymbol{f}_{g1}/\boldsymbol{M}_1 - \boldsymbol{f}_{gi}/\boldsymbol{M}_i)$ in (7) with respect to basis $\mathcal{B}_i$ can be approximated by

$$\begin{aligned} (\boldsymbol{f}_{g1}/\boldsymbol{M}_1 - \boldsymbol{f}_{g1}/\boldsymbol{M}_i)_{ik} &= \omega_{io}^2\{\|\boldsymbol{r}_i\|(\boldsymbol{e}_{ri} \,.\, \boldsymbol{e}_{ik}) - (\|\boldsymbol{r}_i\|/\|\boldsymbol{r}_i\|)^3(\rho_{1ik} + \|\boldsymbol{r}_i\|(\boldsymbol{e}_{ri} \,.\, \boldsymbol{e}_{ik}))\} \\ &\cong -\omega_{io}^2\rho_{1ik}, \quad k = x, y, z, \end{aligned} \tag{9}$$

where $\omega_{io}^2(t) = \mu/\|\boldsymbol{r}_i(t)\|^3$ is the orbital angular speed of the i-th spacecraft about the origin of $\mathcal{F}_o$ at the time t. Substituting (9) into (7), making use of the identity $\boldsymbol{\omega}_i \,\boldsymbol{x}\, (\boldsymbol{\omega}_i \,\boldsymbol{x}\, \boldsymbol{E}_i) = (\boldsymbol{\omega}_i \cdot \boldsymbol{E}_i)\boldsymbol{\omega}_i - \|\boldsymbol{\omega}_i\|^2 \boldsymbol{E}_i$, and assuming $\omega_{io} \cong \omega_o$ (a positive constant) lead to

$$\ddot{\boldsymbol{E}}_i + \dot{\boldsymbol{\omega}}_i \times \boldsymbol{E}_i + 2\boldsymbol{\omega}_i \times \dot{\boldsymbol{E}}_i + (\boldsymbol{\omega}_i \cdot \boldsymbol{E}_i)\boldsymbol{\omega}_i + (\omega_{io}^2 - \|\boldsymbol{\omega}_i\|^2)\boldsymbol{E}_i = \mathcal{D}_i^2(\boldsymbol{h}_i) + \omega_o^2 \boldsymbol{h}_i + \boldsymbol{u}_{c1} - \boldsymbol{u}_{ci}. \tag{10}$$

To derive a formation-keeping control law for the i-th follower, we consider the time rate-of-change of the positive definite function $V_i = (K_{1i}\boldsymbol{E}_i \cdot \boldsymbol{E}_i + \dot{\boldsymbol{E}}_i \cdot \dot{\boldsymbol{E}}_i)/2$, $K_{1i} > 0$. It can be verified [8] that if we set

$$\begin{aligned} \boldsymbol{u}_{ci} = \; & \boldsymbol{I}_i^{-1}(-\boldsymbol{\omega}_i \times (\boldsymbol{I}_i\boldsymbol{\omega}_i) + \boldsymbol{\tau}_{ci}) \times (\boldsymbol{h}_i - \boldsymbol{E}_i) + (\boldsymbol{\omega}_i \cdot (\boldsymbol{h}_i - \boldsymbol{E}_i))\boldsymbol{\omega}_i \\ & + (\omega_{io}^2 - \|\boldsymbol{\omega}_i\|^2)(\boldsymbol{h}_i - \boldsymbol{E}_i) + K_{1i}\boldsymbol{E}_i + K_{2i}\dot{\boldsymbol{E}}_i + 2\boldsymbol{\omega}_i \times \dot{\boldsymbol{h}}_i + \ddot{\boldsymbol{h}}_i + \boldsymbol{u}_{c1}, \end{aligned} \tag{11}$$

where $\dot{\boldsymbol{h}}_i = d\boldsymbol{h}_i/dt_i$, $\ddot{\boldsymbol{h}}_i = d^2\boldsymbol{h}_i/dt_i^2$, and K_{2i} is a positive constant, then $dV_{i1}/dt = -K_{2i}\|\dot{\boldsymbol{E}}_i\|^2 < 0$.

The equation for $[\boldsymbol{E}_i]_i$ corresponding to the feedback-controlled system is given by

$$[\ddot{\boldsymbol{E}}_i]_i + (K_{2i}[\boldsymbol{I}] + 2\boldsymbol{Q}([\boldsymbol{\omega}_i]_i(t)))[\dot{\boldsymbol{E}}_i]_i + K_{1i}[\boldsymbol{E}_i]_i = 0 \tag{12}$$

Since $\boldsymbol{Q}([\boldsymbol{\omega}_i]_i)$ is slew-symmetric, all solutions $[\dot{\boldsymbol{E}}_i, \boldsymbol{E}_i]_i\,(t) \longrightarrow (O, 0) \epsilon \mathcal{R}^6$ as $t + \infty$ for any $K_{1i}, K_{2i} > 0$.

Remarks. RI: The model-dependent Control law (11) corresponds to state-feedback linearization controls which involve partial cancellation of the terms in (10). Assuming perfect cancellation, there is no coupling between the equations for the tracking errors of followers given by (12). Since perfect cancellation is not achievable physically due to inaccurate knowledge of the model parameter values, sensor errors, act uator saturation, and unmodelled external dist urbances, it is of import ante to determine the effect of imperfect cancellation on the behavior of the feedback-controlled system. Here, we model this imperfection by introducing a persistent disturbance N in (12) as follows:

$$\begin{aligned} & [\ddot{\boldsymbol{E}}_i]_i + (K_{2i}[\boldsymbol{I}] + 2\boldsymbol{Q}([\boldsymbol{\omega}_i]_i(t)))[\dot{\boldsymbol{E}}_i]_i + K_{1i}[\boldsymbol{E}_i]_i \\ & = \boldsymbol{N}(t, [\boldsymbol{\omega}_i(t)]_i \cdot [\rho_{ji}(t)]_i, [\boldsymbol{u}_{cj}(t)]_i, [\boldsymbol{E}_i]_i, [\dot{\boldsymbol{E}}_i]_i), \end{aligned} \tag{13}$$

and require that the zero state is totally stable [7]. Since $\boldsymbol{B}_1$ is a skew-symmetric matrix, the zero state of (12) is uniformly asymptotically stable for any $K_{1i}, K_{2i} > 0$. Then, total stability of the zero state follows from a well-known theorem of Malkin [7].

R2: The control laws (11) for the followers require the knowledge of its own attitude control law τ_{ci} and the control law $\boldsymbol{u}_{c1}$ of the leader. The latter information must be transmitted to the follower spacecraft. Note also that control laws (11) can be rewritten as

$$\begin{aligned} \boldsymbol{u}_{ci} = \boldsymbol{I}_i^{-1}(-\boldsymbol{\omega}_i \times (\boldsymbol{I}_i\boldsymbol{\omega}_i) + \boldsymbol{\tau}_{ci}) \times \boldsymbol{E}_i - (\boldsymbol{\omega}_i \cdot \boldsymbol{E}_i)\boldsymbol{\omega}_i - (\omega_{io}^2 - \|\boldsymbol{\omega}_i\|^2)\boldsymbol{E}_i \\ + K_{1i}\boldsymbol{E}_i + K_{2i}\dot{\boldsymbol{E}}_i + \boldsymbol{u}_{ci} + \mathcal{D}_i^2(\boldsymbol{h}_i) + \omega_{io}^2\boldsymbol{h}_i. \end{aligned} \tag{14}$$

The terms $\boldsymbol{u}_{ci}$ and $\mathcal{D}_i^2(\boldsymbol{h}_i) + \omega_{io}^2\boldsymbol{h}_i$ in (11') correspond to a feed-forward control. When the norms of these terms are large, the norm of $\boldsymbol{u}_{ci}$ is also large. This situation may be alleviated by replacing the term by a suitable scaling depending on the norm of $\mathcal{D}_i^2(hi) + \omega_{io}^2\boldsymbol{h}_i$. In the important special case where the spacecraft move in a nearly circular LEO and the deviation vector $\boldsymbol{h}_i$ rotates about the Earth's center with angular velocity $\omega_i \cong \omega_{io}$, then $\mathcal{D}_i^2(hi) + \omega_{io}^2\boldsymbol{h}_i \cong O$ or $\boldsymbol{h}_i$ is close to a solution of the simple harmonic oscillator equation $d^2\boldsymbol{h}_i/dt_o^2 + \omega_{io}^2\boldsymbol{h}_i = O$.

R3: It is evident that if each follower applies control law (11), then the desired formation pattern $\mathcal{P} = \mathrm{P(i)}$, $t \geq 0$, is asymptotically stable, i.e. given any real number $\epsilon > 0$, there exists a $\delta > 0$ such that $\Delta(t) < 6 \Rightarrow \Delta(t) < \epsilon$ for all $t \geq 0$. Moreover, $\Delta(t) \rightarrow O$ as $t \rightarrow \infty$, where

$$\Delta(t) = \left(\sum_{i=2}^{3} \{\sigma_{1i}\|\boldsymbol{E}_i(t)\|^2 + \sigma_{2i}\|\dot{\boldsymbol{E}}_i(t)\|^2\} \right)^{1/2}, \tag{15}$$

and σ_{1i}, σ_{2i} are specified positive weighting coefficients. Here, asymptotic stability is only local in the sense that the convergence of $\Delta(t)$ to O as $t \to 00$ is attained if the deviation of the initial formation pattern at t = O from the desired one is sufficiently small. In physical situations, the possibility of collision between spacecraft must also be considered.

R4: It has been shown recently [9] that under certain mild conditions, asymptotic stability of the formation pattern $\mathcal{P} = \mathcal{P}(t)$ can be achieved using simplified versions of control laws (11') given by

$$\boldsymbol{u}_{ci} = K_{1i}\boldsymbol{E}_i + K_{2i}\dot{\boldsymbol{E}}_i + \{\boldsymbol{u}_{ci} + \mathcal{D}_i^2(\boldsymbol{h}_i) + \omega_{io}^2\boldsymbol{h}_i\}, i = 2,3, \tag{16}$$

which correspond to proportional-plus-rate feedback plus feed-forward controls. The inclusion of the feed-forward control is essential for asymptotic stability.

4 Attitude Control

Let the desired attitude and angular velocity of the i-th follower at time t relative to the inertial coordinate system $\mathcal{F}_o$ be specified respectively by the MCS $\mathcal{F}_i^d(t)$and @(t), which may depend on the attitude and angular velocity of the leader, eg. the desired attitude and angular velocity of the i-th follower correspond exactly to $\mathcal{F}_1(t)$ and $\boldsymbol{\omega}_1(t)$ respectively. It is of interest to control the *relative* attitudes and angular velocities between the spacecraft. Here, we shall derive control laws for the followers which are expressed in terms of their instantaneous attitudes and angular velocities of relative to the inertial coordinate system $\mathcal{F}_o$ or to the MCS $\mathcal{F}_1$.

Let the unit quaternion corresponding to $\mathcal{F}_i^d(t)$ relative to the inertial coordinate system $\mathcal{F}_o$ be denoted by $\boldsymbol{q}_i^d(t) = [\widehat{\boldsymbol{q}}_i^d(t), q_{i4}^d(t)]^T$. We assume that $\boldsymbol{q}_i^d$ and $\boldsymbol{\omega}_i^d$ are consistent in the sense that they satisfy (3) and (4) with" control torque τ_i^d. We introduce the deviations $\delta\boldsymbol{q}_i \overset{\Delta}{=} \boldsymbol{q}_i^d - \boldsymbol{q}_i = [\widehat{\boldsymbol{q}}_i^d - \widehat{\boldsymbol{q}}_i, q_{i4}^d - q_{i4}]^T$ and@ $\overset{\Delta}{=} \boldsymbol{\omega}_i^d - \boldsymbol{\omega}_i$. It can be verified that $\delta\boldsymbol{q}_i$ satisfies

$$\left.\begin{array}{lcl} \frac{d\delta\widehat{\boldsymbol{q}}_i}{dt_o} & - & (q_{i4}\boldsymbol{\omega}_i^d - q_{i4}\boldsymbol{\omega}_i - \boldsymbol{\omega}_i^d \times \widehat{\boldsymbol{q}}_i^d + \boldsymbol{\omega}_i \times \widehat{\boldsymbol{q}}_i)/2, \\ \frac{d\delta q_{i4}}{dt} & = & -(\boldsymbol{\omega}_i^d \cdot \widehat{\boldsymbol{q}}_i^d - \boldsymbol{\omega}_i \cdot \widehat{\boldsymbol{q}}_i)/2. \end{array}\right\} \tag{17}$$

To derive an attitude control law, we consider the following positive definite function $V_{1i} = K_{qi}\dot{V}_{1i} + \ddot{V}_{1i}$ defined on R^7, where K_{qi} is a given positive constant and $\dot{V}_{1i} = \delta q_{i4}^2 + \delta\widehat{\boldsymbol{q}}_i \cdot \delta\widehat{\boldsymbol{q}}_i$; $\ddot{V}_{1i} = (\delta\boldsymbol{\omega}_i \cdot I_i\delta\boldsymbol{\omega}_i)/2$. The time derivative of V_{1i} along the solutions of the equations for $\delta q_i, \delta\omega_i$ is given by

$$\frac{dV_{1i}}{dt} = \{K_{qi}(q_{i4}^d\delta\widehat{\boldsymbol{q}}_i - \delta q_{i4}\widehat{\boldsymbol{q}}_i^d - \widehat{\boldsymbol{q}}_i^d \times \delta\widehat{\boldsymbol{q}}_i) + \tau_{ci}^d - \tau_{ci} - \boldsymbol{\omega}_i^d \times (\boldsymbol{I}_i\delta\boldsymbol{\omega}_i)/2\} \cdot \delta\boldsymbol{\omega}_i. \tag{18}$$

Thus, if we set

$$\tau_{ci} = K_{qi}(q_{i4}^d\delta\widehat{\boldsymbol{q}}_i - \delta q_{i4}\widehat{\boldsymbol{q}}_i^d - \widehat{\boldsymbol{q}}_i^d \times \delta\widehat{q}_i) + \tau_{ci}^d - \boldsymbol{\omega}_i^d \times (\boldsymbol{I}_i\delta\boldsymbol{\omega}_i)/2 + K_{\omega i}I_i\delta\boldsymbol{\omega}_i, \tag{19}$$

where $K_{\omega i}$ is a positive constant, then $dV_{1i}/dt = -K_{\omega i}\delta\boldsymbol{\omega}_i \cdot I_i\delta\boldsymbol{\omega}_i \leq 0$ and $Vii(t) \leq V_{1i}(0)$ for all $t \geq 0$ implying uniform boundedness of $\|\delta\omega_i(t)\|$ for all $t \geq 0$, By considering d^2V_{1i}/dt^2, and making use of Barbalat's Lemma [10], we can deduce that $(dV_{1i}/dt)(t) \to O$ as $t \to 00$, or $\omega_i(t) \to \omega_i^d(t)$ as $t \to \infty$. But it does not follow that $\delta q_{i4}(t) \to O$ and $\delta\widehat{\boldsymbol{q}}_i(t) \to$ O as $t \to 00$.

To proceed further, we make use of the fact that the quaternion $(\Delta\hat{q}_i, \Delta q_{i4})$ of the desired attitude or $\mathcal{F}_i^d$ *relative to* $\mathcal{F}_i$ is related to the quaternions $(\hat{q}_i^d, q_{i4}^d)$ for $\mathcal{F}_i^d$, and $(\hat{q}_i, q_{i4})$ for $\mathcal{F}_i$ relative to the inertial coordinate system $\mathcal{F}_o$ by

$$\Delta\hat{q}_i = q_{i4}\hat{q}_i^d - q_{i4}^d\hat{q}_i - \hat{q}_i \times \hat{q}_i^d, \quad \Delta q_{i4} = q_{i4}q_{i4}^d + \hat{q}_i \cdot \hat{q}_i^d. \tag{20}$$

When $\mathcal{F}_i$ coincides with $\mathcal{F}_i^d$ (i.e. $\delta\hat{q}_i = O$ and $\delta q_{i4} = O$), we have $\Delta\hat{q}_i = O$ and $\Delta q_{i4} = 1$. Using (18), control law (17) can be rewritten as

$$\tau_{ci} = -K_{qi} \Delta q_i + \tau_{ci}^{\star} - \omega_i^d X (I_i\delta\omega_i)/2 + K_{\omega i}I_i\delta\omega_i. \tag{21}$$

Following an analysis similar to that given in [II], conclude that $\Delta\hat{q}_i$ and $\delta\omega_i(t) \to 0$ as $t \to \infty$ for any positive K_{qi} and $K_{\omega i}$.

5 Implementation of Control Laws

We observe that the implementation of control laws (11) and (19) for formation keeping requires a knowledge of $[E_i]_i, [\dot{E}_i]_i, [\omega_i]_i; [\tau_{ci}]_i$ and $[u_{ci}]_i$ at any time t. The quantities $[E_i]_i$ and $[\dot{E}_i]_i$ can be determined from $[\rho_{1i}]_i$ and $[\dot{\rho}_{1i}]_i$ which require measurement of the position and velocity of the leader relative to the i-th follower spacecraft. These quantities can also be obtained by transmitting the position and velocity of leader to the i-th follower. Also, the control of the leader at any time must also be transmitted to the i-t h follower spacecraft. When one or more spacecraft failure occurs, one may adopt the following backup schemes for control law implement aiton depending on the nature of failure:

(i) Inter-spacecraft Communicaiton System Failure: One may obtain estimates of $[\rho_{1i}]_i$ and $[\dot{\rho}_{1i}]_i$ by using on-board optical range sensors, or by setting the relative position and velocity between the failed and active spacecraft at their nominal values temporarily until the failure is recovered.

(ii) Overall Spacecraft Failure: Here the failure is sufficiently severe such that the failed spacecraft is no longer useful. In this case, it should be removed from the formation by deorbiting or by manual retrieval, If the failed spacecraft is replaced by a backup spacecraft, then it is necessary to reconfigure the formation, The control laws for steering the remaining active spacecraft from the old to the new formation requires separate consideration. This aspect will be discussed elsewhere.

We note also that in the derivation of foregoing control laws, no constraints have been imposed on the magnitude of the control variables. In the presence of bounded controls, one expects that the rate of decay of $\|([E_i]_i, [\dot{E}_i]_i)(t)\|$ and $\|(\delta\omega_i, \delta q_i)(t)\|$ to zero would be reduced by when one or more of the control variables takes on its extreme values.

Finally, for a real mission, it is necessary to consider discrete-time versions of the proposed control laws. In view of the limited fuel on-board, it is generally undesirable to have continuously acting controls. Therefore, the system response corresponding to the cent rol laws derived here serves as a basis for comparison between the idealized and the actual responses.

6 Fleet Coordination

For a fleet of spacecraft, one may require complete autonomy in each spacecraft in the sense that all the decisions for determining its future behavior are made on-board without the assistance of external agents. Although this approach provides enhanced operational reliability, it may not be cost effective since each spacecraft must contain all the essential hardware and software for

coordination and control. An alternative approach is to require each spacecraft to have only the basic hardware and software for attitude control and orbital nameuvering. The more complex tasks in fleet coordination and control are shared by all the spacecraft in the fleet. Moreover, some of the spacecraft may be equipped with special hardware and software to perform particular tasks for the entire fleet.

The fleet coordination is achieved with the aid of an inter-spacecraft communication network (eg. radio or optical links). This network has the following basic functions:

(i) Communicating the necessary data for fleet formation-keeping and relative attitude control;

(ii) Linking the computers in the spacecraft to form a distributed computing network thereby increasing the computational capability of the fleet for more computational intensive tasks such as on-board interferometer data processing,

In the realization of the first function, each fleet leader broadcasts its position and attitude with respect to a specified inertial frame, and the follower spacecraft broadcast their positions and velocities relative to their leader to achieve formation alignment.

7 Simulation Studies

Extensive simulation studies have been made to determine the performance of the proposed control laws for formation keeping and attitude regulation in the presence of actuator saturation, variations in spacecraft parameters, and loss of communication between spacecraft. Only typical results will be presented here.

We assume that the leader of the spacecraft triad moves along an inclined circular orbit $\mathcal{O}_1$ about the Earth with inclination angle $(\pi/2 - \varphi_{inc})$ and ascending node along the Y-axis. For convenience, we introduce a geocentric fixed cartesian coordinate frame $\mathcal{F}_o$ with origin O at the Eart h's center along with a spherical coordinate system (r, θ, ϕ) wit h orthonormal basis $\{e_r, e_\theta, e_\phi\}$. Its motion in spherical coordinates is given by

$$\begin{aligned} r_1(t) &= r_o, \quad \theta_1(t) = cos^{-1}\{cos(\varphi_{inc})\,cos(\theta_o - \omega_o t)\}, \\ \phi_1(t) &= tan^{-1}\{-tan(\theta_o - \omega_o t)/sin(\varphi_{inc})\}, \end{aligned} \tag{22}$$

where r_o is a given orbital radius, and $\omega_o = \sqrt{\mu/r_o^3}$. Here, for simplicity, we have set the desired orbit al radius for all spacecraft to r_o.

The desired motions for the second and third spacecraft correspond to two circular orbits with the same inclination angle $(\mu/2 - \varphi_{inc} = 8.2\mu/180$ rad., but with ascending nodes at $(r, \theta, \phi) = (r_o, \mu/2, \Delta\phi)$ and $(r_o, \mu/2, -\Delta\phi)$ respectively, where $\Delta\phi = 9.5\mu/180$ rad. We adopt the simplified control law (11") for formation keeping, where the deviation vector $h_2(t)$ is given by

$$\begin{aligned} h_2(t) \;=\; & r_o\{(-cos(\Delta\phi)sin\,(\varphi_{inc})cos(\bar{\theta}(t)) + sin(\Delta\phi)sin(\bar{\theta}(t)))e_X \\ & + (sin(\Delta\phi)sin(\varphi_{inc})\,cos(\bar{\theta}(t)) + cos(\Delta\theta)sin(\bar{\theta}(t)) - sin(\theta_o - \omega_o t))e_Y \\ & + (cos(\varphi_{inc})cos(\bar{\theta}(t) - \Delta\theta) - cos(\theta_o - \omega_o t))e_Z. \end{aligned} \tag{23}$$

where $\bar{\theta}(t) = 19_0 - \omega_o t - \Delta\theta$. The deviation vector $h_3(t)$ has the same form as (21) except with $\Delta\phi$ replaced by $-\Delta\phi$. Evidently, hi(t) satisfies $d^2 h_i\,(t)/dt_o^2 + \omega_o^2 h_i(t) = O$ for all t and i = 2,3. To specify the desired attitude of the i-th follower, we introduce the (1-2-3) Euler angles $(\Theta_i, \Psi_i, \Phi_i)$ corresponding to a rotation of Θ_i about the X-axis followed by a rotation of Ψ_i about the rotated

Y-axis, and a rotation of Φ_i about the rotated Z-axis. The desired Euler angles for the followers are given by:

$$\begin{aligned}\Theta_2^d(t) &= \Theta_3^d(t) = \bar{\theta}(t),\ \Psi_2^d(t) = \Psi_3^d(t) = \varphi_{inc},\\ \Phi_2^d(t) &= \Delta\phi,\ \Phi_3^d(t) = -\Delta\phi.\end{aligned} \tag{24}$$

Thus, the desired attitude of the i-th spacecraft can be expressed in terms of the following quaternions:

$$\begin{aligned}q_{24}^d(t) &= \{1 + cos(\Delta\phi)cos(\varphi_{inc}) - sin(\Delta\psi)sin(\varphi_{inc})\, sin(\bar{\theta}(t))\\ &\quad + cos(\bar{\theta}(t))(cos(\Delta\psi) + cos(\varphi_{inc}))\}^{\frac{1}{2}}\ /2; \qquad (25)\\ q_{21}^d\ (t) &= \{sin(\Delta\phi)sin(\varphi_{inc})cos(\bar{\theta}(t)) + sin(\bar{\theta}(t))(cos(\Delta\phi)\\ &\quad + cos(\varphi_{inc}))\}/(4q_{24}^d(t))\ ; \qquad (26)\\ q_{22}^d(t) &= \{sin(\varphi_{inc})(1 + cos(\Delta\phi)cos(\bar{\theta}(t))) - sin(\Delta\phi)sin\ (\bar{\theta}(t))\}/(4q_{24}^d(t)); \qquad (27)\\ q_{23}^d(t) &= \{cos(\Delta\phi)\ sin(\varphi_{inc})sin(\bar{\theta}(t))\\ &\quad + sin(\Delta\phi)cos(\bar{\theta}(t)) + cos(\varphi_{inc}))\}/(4q_{24}^d(t)). \qquad (28)\end{aligned}$$

The quaternions corresponding to the desired attitude of the third follower have the same form as (23) except with $\Delta\phi$ replaced by – A+. We require every spacecraft to spin about its z-axis with constant angular speed ω_s. Thus, the desired angular velocity for the i-th follower is given by

$$\omega_i^d = \omega_o cos(\varphi_{inc})e_X + \omega_o sin(\varphi_{inc})e_Z + \omega_s e_{iz}^d = \omega_s e_{ix}^d + \omega_s e_{iz}^d, \tag{29}$$

where $\{ e_{ix}^d, e_{iy}^d, e_{iz}^d \}$ corresponds to the basis of the body coordinate system $\mathcal{F}_i^d$ associated with the i-th follower with the desired attitude.

Figure 4 shows a typical time-domain response of the MS fleet with the simplified formation-keeping control law(11") and attitude control law (19) in the presence of actuator saturation, The spacecraft parameter values used in the simulation study are given in Table 1. The corresponding time-domain response of the MS fleet with the $[u_{c1}]_i$ term in (11") set to zero (to simulate the loss of communication between the MS) was also determined. The results do no differ significantly from those shown in Fig. 4. Next, the effect of inertia pertrubations on the time-domain response of the spacecraft triad was studied. It was found that the qualitative behavior of the response is essentially identical to that of the unperturbed case.

8 Concluding Remarks

In this paper, control laws for a spacecraft triad moving in formation have been derived using a simplified model for a rigid spacecraft. These control laws require the knowledge of the relative displacements and attitudes of the spacecraft and its neighbors. Simulation results based on a generic spacecraft model showed that the derived control law are effective in formation and relative attitude alignment provided that the magnitude of the initial deviation from the desired state is sufficiently small so that collisions between the spacecraft do not occur. Finally, in this work, important factors such as data processing time-delay and time discretization arising in physical implementation have not been taken into consideration. Nevertheless, the results reveal the basic structure of the control laws and the required inter-spacecraft data required for their implementation. Finally, the problems associated with the physical implementation of the control laws in terms of the state-of-the-art hardware and fuel consumption for control are not considered here, and they require further study.

M_i	mass of MS	$10 kg.$
I_{ix}	moment of inertia about x-axis	$0.3646\, kgm^2.$
I_{iy}	moment of inertia about y-axis	$0.2734 kgm^2.$
I_{iz}	moment of inertia about z-axis	$0.3125\, kgm^2.$
r_o	desired orbital radius of MS	$7.13814\ x10^6 m.$
$\omega_o = \sqrt{\mu/r_o^3}$	orbital angular speed of fleet leader	$0.001\, rad./sec.$
ω_s	desired spin speed about z-axis	$0.01 rad./sec.$
$\varphi_{inc} = \pi/2$	inclination angle of reference orbits	$8.2\pi/180 rad.$
$\triangle\phi$	azimuthal angle associated with the ascending node of reference orbits	$0.2 rad.$
$\triangle\theta$	MS separation angle	$\pi/120 rad.$

Table 1: Values of microspacecraft and orbital parameters for simulation study,

Acknowledgement

This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronaut ics and Space Administration.

References

[1] Statchnik, R., Ashlin, K. H., *Bulletin of American Astronomical Sot.* Vol. 16, No. 3, 1984, pp. 818-827.

[2] Statchnik, R., Melroy, R., and Arnold, D., "Multiple Spacecraft Michelson Stellar Interferometry" *Proc. SPIE, Instrumentation in Astronomy V* Vol. 445, pp. 358-369, 1984.

[3] Johnson, M. D. and Neck, K. T., "Free Flyer Optical Interferometry Trajectory Analysis" JPL 4/30-5/2, 1990.

[4] DeCue, A. B., "Multiple Spacecraft Optical Interferometry, Preliminary Feasibility Assessment" JPL Technical Report D-1811, Aug. 1991.

[5] Wertz, J. R. (Ed.), *Spacecraft Attitude Determination and Control* D. Reidel Pub. Co., Boston, 1980.

[6] Wang, P. K. C. "Navigation Strategies for Multiple Autonomous Mobile Robots Moving in Formation" *J. Robotic Systems* Vol. 8, pp. 177-195, 1991.

[7] Hahn, W., *Stability of Motion* Springer-Verlag, New York, 1967.

[8] Wang, P. K. C. and F. Y. Hadaegh, "Coordination and Control of Multiple Micro-spacecraft Moving in Formation" UCLA Engr. Rpt. 94-102, Sept. 1994,

[9] Wang, P. K. C. and F. Y. Hadaegh, "Simple Formation-keeping Control Laws for Multiple Microspacecraft" UCLA Engr. Rpt. 95-130, August, 1995.

[10] Popov, V. M., *Hyperstability of Automatic Control Systems* Springer Verlag, New York, 1973.

[11] Wen, J. T. and Kreutz-Delgado, K., "The Attitude Control Problem" *IEEE Trans. Auto. Control* Vol. 36, No. 10, Oct. 1991, pp. 1148-1162

List of Figures

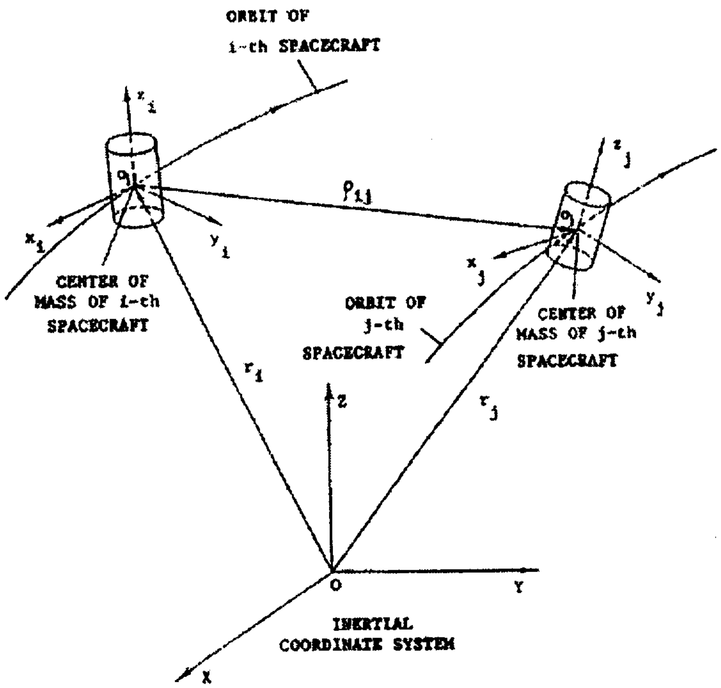


Figure 1. Sketch of inertial and moving coordinate systems.

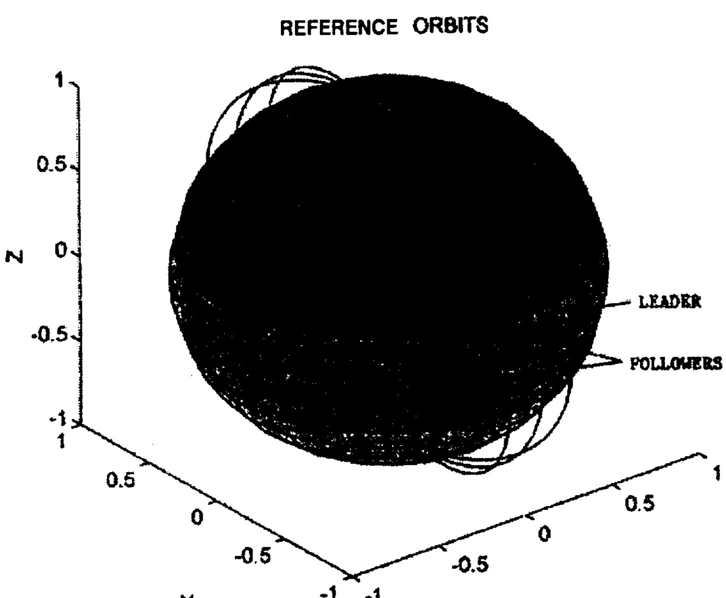


Figure 2. Exaggerated sketch of the reference orbits of spacecraft triad moving in formation along LEO about the Earth.

Figure 3. Time-domain response of the spacecraft triad with simplified formation-keeping control law (11") and attitude control law (19) with K_{1i}=0.5, K_{2i}=2.0, $K_{\omega i}$=1.5, K_{qi}=0.4, 1=2,3, in the presence of actuator saturation, and with initial states:

$[E_2(0), \dot{E}_2(0)]_2$ = [52 -5, 00 O];

(CO,(o)), = (0.02 0.02 0.02);

$q_2(0)$= [0.3 0.10.20.927261 8];

$[E_3(0), \dot{E}_3(0)]_1$ = [-1 2-1, 0 0 0];

$[\omega_3(0)]_3$ = [-0.01 0.015- 0.01];

$q_3(0)$ = [-0.2 0.20.3 0.910433];

The Follower spacecraft 2 is shown with solid lines and the follower spacecraft 3 is shown with dashed lines; Saturation levels:

$|f_{cij}| \leq 1$ N; $|\tau_{cij}| \leq 0.05$ N.m, 1=2,3; j=x,y,z.

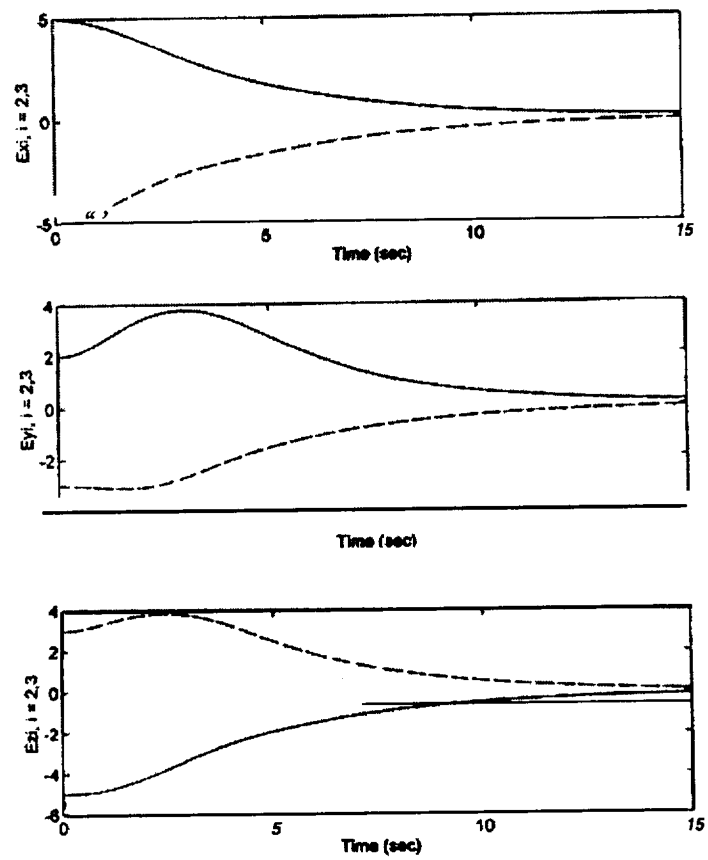


Figure 3a. Positional tracking errors (m) *vs.* time.

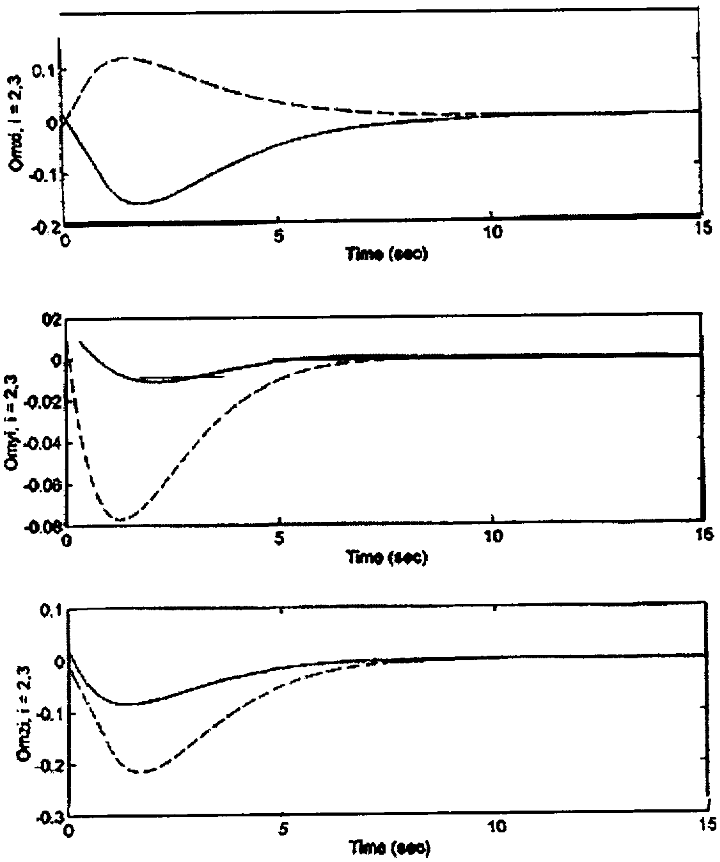


Figure 3b. Angular velocities (m/see) vs. time.

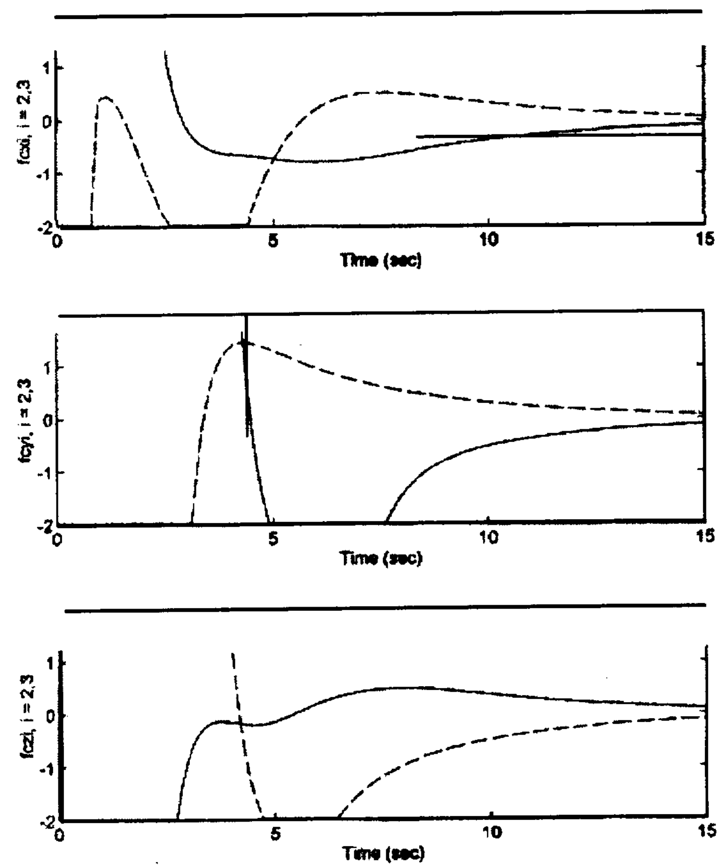


Figure 3d. Control forces (N) vs. time.

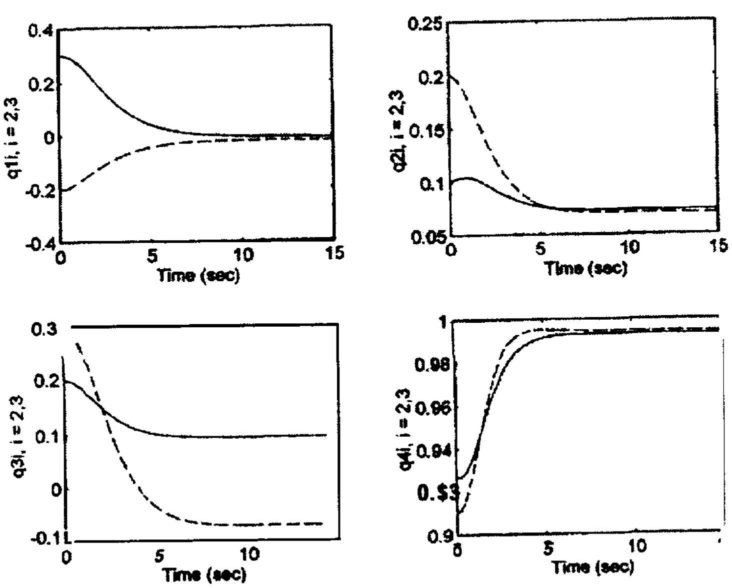


Figure 3c. Quaternions vs. time.

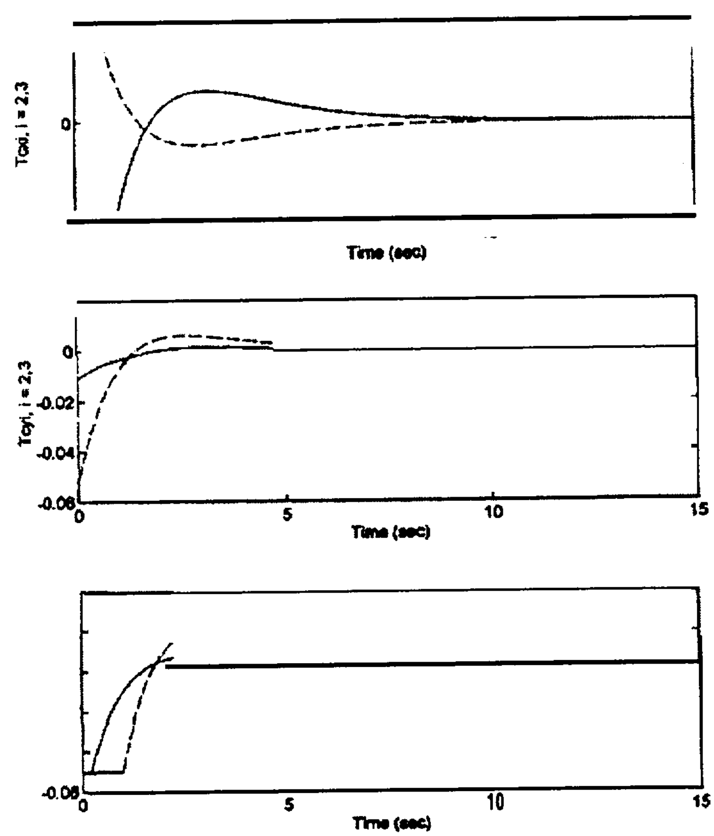


Figure 3e. Control torques (N.m) vs. time.

URC97162

Multi-window Controllers for Autonomous Space Systems

B. J. Lurie and F. Y. Hadaegh
Jet Propulsion Laboratory
California Institute of technology
Pasadena, CA 91109

Abstract.

Multi-window controllers *[1,2]* select between elementary linear controllers using nonlinear windows based on the amplitude and frequency content of the feedback error. The controllers are relatively simple to implement and perform much better than linear controllers. The commanders for such controllers only order the destination point and are freed from generating the command time-profiles.

The robotic missions rely heavily on the tasks of acquisition and tracking. For autonomous and optimal control of the spacecraft, the control bandwidth must be larger while the feedback can (and, therefore, must) be reduced.. Combining linear compensators via multi-window nonlinear summer guarantees minimum phase character of the combined transfer function. It is shown that the solution may require using several parallel branches and windows, Several examples of multi-window nonlinear controller applications are presented.

1. Composite nonlinear controllers

R was demonstrated in [1-5] that some nonlinear controllers perform better than any linear controllers. Therefore, the optimal controller is, generally, nonlinear.

For a sufficiently small region in the state space, the nonlinear optimal controller can be approximated well by a linear controller. For the adjacent region, another linear controller can be designed that would be optimal over this region, and so on. Then a cental design problem arises as of how to integrate these locally-optimal linear controllers into a composite nonlinear controller, and in particular, how to provide smooth transitions between these linear laws when the time trajectory crosses the borders between these regions.

The transitions between the control modes can be defined by ***participation rules*** illustrated in Fig. 1. Over some transition interval of a variable or condition, the controller action is the sum of the actions of the adjacent regional control modes, and at the ends of the interval only one of the modes (or parameters, or actions) takes place. The linear transition rule Fig. l(a) can be expressed as

$$action = action_1(1 - k) + k \cdot action_2, \qquad (1)$$

where the scalar k *is* changing from O to 1 over the transition interval. A smoother rule of the modes' changing is illustrated in Fig. l(b). Commonly, the precise shape of the participation rules does not matter much as long as it is monotonic, not too steep, and not too shallow.

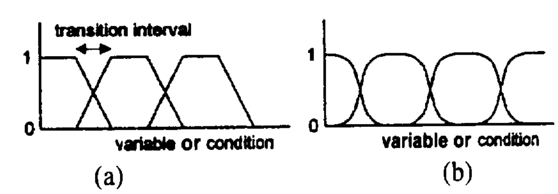


Fig. 1. Participation functions in composite controllers

The monotonic shape of the participation rules does not yet guarantee the smoothness of the transition. R is also required that the adjacent control laws mix well, i.e., the combined action of the adjacent controllers exceeds that of each individual controller. This is not always the case. For example, even if some residue that needs to be cleaned out can be removed by either an acid or a base, an acid and a base should not be used as a mixture with gradually changing content. For regulation of a reactive electrical current, a variable capacitor or a variable inductor can be used, but these elements should not be combined in series or parallel since they might produce resonances. A low-

pass link should not be mixed in parallel with a high-pros link or else notches and n.m.p. shift might result [2,5]. ***Fuzzy logic controllers*** break each smooth transition into several discrete steps. This increases the total number of regions with different control laws. Since these regions become small, fuzzy logic control can use low order regional control laws. Hence, fuzzy control design can be based on phase-plane partitions, and on passivity theory. However, in fuzzy logic controllers many variables need to be sensed and processed to define the boundaries of the regions.

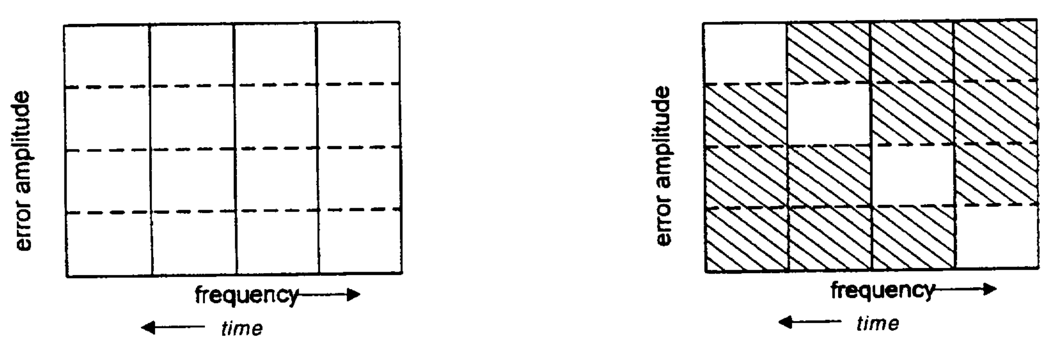
What region size is optimal for composite controllers? There are two advantages of making the regions small. The first is that the control laws in the adjacent regions might become very similar which enables smooth transitions between them without taking special precautions. The second advantage is, because the linear controller can be of low order, the phase plane can be used for the controller analysis and design, and, as claimed by some of fuzzy logic advocates, the controllers can be designed even by those ignorant of control theory. However, when the number of the regions is large, the number of boundaries between them becomes very large. Correspondingly, the number of the decision making algorithms and instruments for changing between the control modes becomes very large. This complicates both the controller design and the designed controller.

On the other hand, higher order linear control laws can be made to remain nearly optimal over a much broader region than low-order laws. This reduces the number of the regions and the number of the boundaries between them thus making the designed controller much simpler. However, for the design of higher order regional control laws, not the phase plane but frequency domain methods should be used. The partition between the regions should be also somehow defined in the frequency domain. This approach requires caution and application of certain rules discussed in Chapter 4 to provide good blending of the regional control laws at the boundaries between the regions. Nevertheless, this approach is not difficult and leads to economical and nearly optimal controllers.

2. Multi-window control

In the following, we will consider a subset of the nonlinear controllers where the control law is dependent only on the amplitude of the error at the output of the feedback summer. The controller is composed of linear operators and non-dynamic (static) nonlinear functions determined by participation rules. It will be shown in examples and by ***some*** argumentation that such controllers provide nearly optimal performance for a wide variety of practical problems.

The sets of the error signals can be bounded by two-dimensional ***windows*** shown in Fig. 2(a). The windows divide the frequency spectrum (or, equivalently, time-response behavior) and the amplitude range. Within each window, some optimal linear operator (probably, of high order) is employed.



(a) (b)

Fig. 2. The multi-window control concept: (a) the choice of the linear controller defined by the error amplitude and frequency content, (b) the linear controller is defined only by the error amplitude

When the amplitudes change, other windows become active and the control law changes. The amplitudes of different frequency components of the error define the composite control law.

The multiwindow nonlinear controller can be implemented as follows: the error is partitioned into components belonging to several windows, the components which belong to a window processed by the window linear operator, and the results combined by nonlinear functions implementing the appropriate participation law. This architecture is referred to as ***multi-window***, and a great number of useful nonlinear control schemes can be cast in this form.

In many practical cases, the correlation between the error's amplitude and the frequency content allows using only the diagonal windows as shown in Fig. 2(b). (The error amplitude can be linearly weighted.) The optimal linear controllers within each window can be designed using Bode integrals as the realizability limitation and the criterion for minimum phase behavior as the condition of smooth blending with the adjacent regions differing in frequency.

High-order linear regional controllers can be made nearly optimal not only within rather wide assigned windows, but also over some reasonably wide margins overlapping the adjacent windows, In this case, with relatively steep participation rules and narrow transition areas as illustrated in Fig, 3, the participation rules and the position of the borders between the windows become not critical.

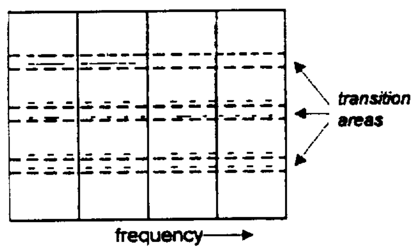


Fig. 3. Transition areas at the window borders

When the transition area is narrow (which often is the case), the transition is only required to be smooth. When the transition area is comparable to the area of the window, the transitional nonlinear control law must also be optimal. The transitional operators can be analyzed and synthesized using either the absolute stability approach or the describing function approach, using as the design constraints the Bode integrals.

The static nonlinearities used to implement the transition between the control modes lend themselves to realization using fuzzy rules. The nonlinearities are chosen (like saturation) such that sharp discontinuities are avoided.

Two-window compensators are widely employed (although not always optimized) in practical feedback systems, particularly, in anti-windup schemes, in acquisition and tracking systems, and in Nyquist-stable systems for provision of global and process stability.

The multiwindow compensator is a nonlinear dynamic compensator. And, vice versa, the NDC discussed in [2], whether they are made as a combination of parallel channels or as links with nonlinear local feedback, are multiwindow controllers. In [2], composite nonlinear controllers are studied as the means of providing global stability and process stability. In this paper, we concentrate on applications.

3. Windup, and anti-windup controllers

The overshoot in a system with saturation for the large amplitude input step can be excessive and persistent – this phenomenon is called ***windup. The*** windup can be many times longer than the overshoot in the linear mode of operation. Windup is caused by a combination of two factors: the error integration in the compensator and the actuator saturation. The saturation limits the return signal and therefore prevents the error signal accumulated in the compensator integrator from being compensated. During the initial period after the step command is applied, when the output is still low and the error is large, the compensator integrates the error. When the time comes at which this integrated error would be compensated by the return signal in the linear mode of operation, this does not happen for large commands, since the return signal is reduced by the actuator saturation. Then, it might take a long time for the feedback to compensate the integrated error. The error "hangs up", and only after some time, the output signal drops to the steady state value.

Using the DF concept, the qualitative explanation for the windup phenomenon goes as follows. Actuator saturation reduces the describing function loop gain thus shifting the equivalent crossover frequency down in frequency. The resulting overshoot is long, corresponding to this low crossover frequency. The value of the windup depends on the loop phase lag. When the phase stability margin is more than 70°, the windup is practically nonexistent, but it is large when the stability margin is 30° or smaller. The windup can be reduced or eliminated by employing nonlinear dynamic compensation.

Windup in *PID* system is commonly reduced or eliminated by placing a saturation in front or after the integrator, or resetting the integrator. A saturation link with different threshold is sometimes placed as well in front of P-term as shown in Fig. 4 thus making a three-window controller.

The diagram in Fig. 2 is somewhat ambiguous since it does not indicate whether the frequency selection or the amplitude selection is performed frost. Often a particular order is required. This order is different in the block diagrams in Fig. 5 which exemplifies two types of architectures for multi-window compensators.

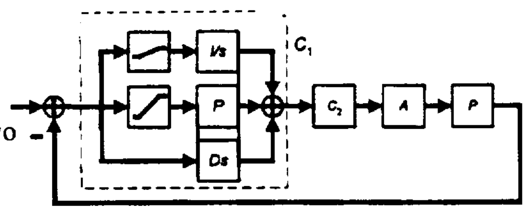
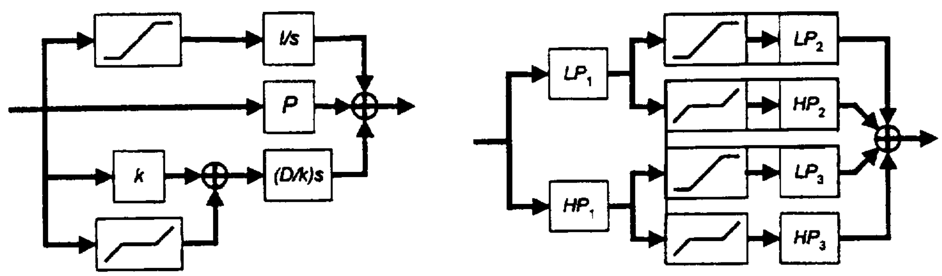


Fig. 4. Saturation in front of *I*- and P-paths.



(a) (b)

Fig. 5. Multi-window compensators with parallel channels

Fig. 5(a) shows compensators with parallel channels. In such compensators, saturation links are commonly placed in the low-frequency channel since this channel's gain response dominates at lower frequencies. At large signal levels, the saturation link reduces the low-frequency gain and the phase lag of the compensator decreases thus reducing or eliminating windup. Placing a saturation link in front of the Z-path in *PID* controller commonly allows making small both the length and the height of the overshoot. In doing so, the value of the saturation threshold is not very critical. Placing a dead-zone element in front of the high-frequency channel with $k < 1$, also reduces the phase lag at large signal amplitudes which helps to eliminate the windup and improve the transient response.

Also, the windows can be placed in the local feedback path of an amplifier in the compensator. The performance of such multiwindow controllers is at least as good as that of the parallel paths controllers, and in this case the design can be placed on firm foundation of Popov absolute stability criterion [2,3,5],

4. Acquisition and tracking

Acquisition and tracking systems, like those used in homing missiles, are designed to operate in two modes: acquisition mode when the error is large, and tracking mode when the error is small. An example of the acquisition/tracking type is a pointing control system for a spacecraft-mounted camera, in which a rapid retargeting maneuver is followed by a slow precise scanning pattern to form a mosaic image. Another example is clock acquisition in the phase-locked loops of telecommunication systems and frequency synthesizers. When the error signal is large, i.e. the system is in the acquisition regime, the controller should respond as rapidly as possible, i.e. the feedback bandwidth should be wide. In the acquisition mode it is not necessary however that the feedback be very large, since the error is big anyway. In contrast, in the tracking regime, the feedback bandwidth needs to be reduced to reduce the output effects of the sensor noise, but the value of the feedback should be made rather large to minimize the tracking error. The differing loop frequency responses for the two modes are depicted in Fig. 6.

While the determination of theoptimal frequency responses for the acquisition mode and for the tracking mode is straightforward, guaranteeing smooth transient responses during transition from acquisition to tracking is not trivial. The transition can generate large transients in the output and error signals. If the transients are excessively large, the target can be de-acquired.

The transition between the responses can be done by switching or, better, by using nonlinear windows: the small errors are directed to the tracking compensator, and the large errors directed into the acquisition compensator. When using the windows, special care must be taken to ensure that all intermediate combined frequency responses of the parallel channels are acceptable. Intermediate response might result in an unstable system, or in a system with small stability margins and, therefore, producing large amplitude transient responses.

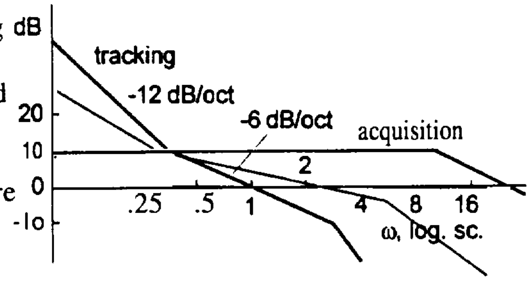


Fig. 6. Acquisition/tracking loop

As an example, let the total loop response be the weighted sum of the acquisition and tracking responses:

$$T = (1-k)T_{acq} + kT_{tr}, \qquad (2)$$

and suppose that k smoothly varies from O to 1 as the transition from acquisition mode to track mode occurs. For a certain value of k, *the* gains in the two paths are equal at the frequency f_1 indicated in Fig. 6. At this point the difference in phase between the two transfer fictions exceeds 180°, and the result is that a zero of the total transfer function T moves into the right half-plane ofs. The transient generated while the system remains in this state can be big and disruptive, even causing the target to be lost.

The conditions for the two parallel path transfer function $W_1 + W_2$ to become n.m.p. when each of the channels is m.p., is given in [2]. When the response is changing between two responses that, which combined, form n.p. lag that causes oscillation in the feedback loop – what happened? Qualitatively, we can describe the effect as follows. When one response dies down and the second response gradually rises to power, there is a time interval when they both are active, with the describing function k producing n.p. loop transfer describing function. When this time interval is long enough, a violent transient can occur that might lead to the target de-acquisition. For example, when the system in Fig. 5 with responses of Fig. 6 was modified so that the compensators are not switched instantly but their outputs combined via nonlinear windows, the overshoot in computer simulation reached 500%.

Therefore, when using two-window controllers with nonlinear blending of the linear controllers, the two compensator responses should not differ as much as those shown in Fig. 6. Hence, the two-window controller, although substantially better than a linear controller, still do not allow implementation of the best possible responses for acquisition and for tracking. This can be done with a three window controller using an intermediate frequency response during the transition which will look like the dotted line in Fig. 6.

5. Time-optimal control

Time-optimal control means changing the output variable between the commanded limits in minimum time, with limited force or power amplitude. For example, shifting a , mass with limited force amplitude in minimum time results in the force profile shown in Fig. 7. This control switches the force from positive to negative values at appropriate instants. The switching must occur at exactly right moments or else the final error will be large. When the plant is uncertain, the moments cannot be exactly calculated, and the open loop control entails considerable errors.

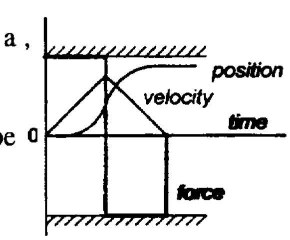


Fig. 7. Time-optimal control of a rigid plant position

When the time-optimal control has to be implemented closed loop-wise, it still should provide swift transitions from the maximum positive to maximum negative values, i.e. transitions very close to switching. The switching controller is a relay controller, i.e. a controller with infinite gain in the loop. An approximation to this controller is a controller with saturation and very large gain of linear links of the loop. The problem of designing such controller is therefore a feedback maximization problem while providing global stability and windup elimination. The solution to this problem also requires using an NDC, i.e. a multi-window controller as was shown in [4,5].

A reasonable complexity linear controller cannot provide time-optimal control. But a nonlinear, even a simple two-window controller, can be quite good. For most common problems, two windows suffice. However, when the transition between the modes must be very fast, or when the dynamic range is large, more windows may be necessary. For example, time-optimal precision pointing of space optical telecommunication systems with a huge ***range*** of error from acquisition to tracking would require more than two windows.

Example 1: Despin Control for S/C Booster Separation. The spacecraft booster is stabilized by spinning at 85 RPM. After separation from the booster, the spacecraft shown in Fig. 8(a) is despun by yo-yo to about 2 RPM, (The yo-yo is a weight at the end of a cable wrapped several times around the spacecraft. When the spacecraft is released from the booster rocket, the weight is also released and begins unwrapping the cable. When all the cable length is unwrapped, the cable is separated from the spacecraft, and the yo-yo takes away most of the rotation momentum.) The remaining spacecraft spin needs to be removed by fining thrusters. The spacecraft is unstable for spinning about z-axis since it is prolate, and the despin should be fast. Because of huge uncertainty in the initial conditions after the separation, with various positions and spin rates and different types of coupling between the axes, the controller design for the despin function must be made very robust, and at the same time, it must perform in a nearly time-optimal fashion. After the despin is complete, the controller must be changed to provide better control accuracy in the cruise mode.

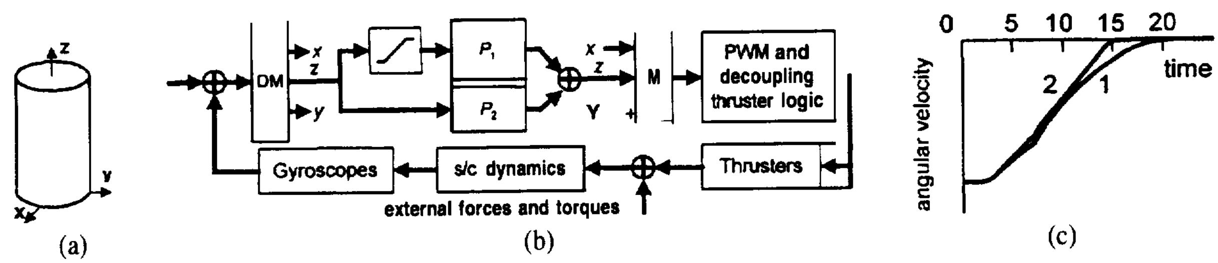


Fig. 8. A spacecraft (a) local frame coordinates, (b) attitude control system block diagram, (c) time-response of z-axis despinning

The controller shown in Fig. 8(b) uses pulse width modulated (PWM) thrusters. Since each thruster produces x-, y-, and z-torques, they are combined in pairs and decoupled by the thruster logic matrix. This renders the control of each axis independent to a certain extent. The problem is however complicated by coupling between the x, y, and z-rotations due to the spacecraft dynamics, including spinning of fuel and oxidizer, initially at the rate of the booster. Due to large plant uncertainty, despin was chosen to be proportional, providing large phase stability margin over the entire frequency range of possible plant uncertainty and x-, y-, and z-controllers coupling.

In the block-diagram, the demultiplexer DM separates vector into its components. The multiplexer M is doing the opposite. The compensators are independent for the x-, y-, and z-rotations, i.e., the controller matrix is diagonal.

When the controllers' gains were chosen such as to despin the s/c without substantial overshoot, the z-axis response was as shown in Fig, 8(c), curve 1. It is seen that the control is not time-optimal.

A better controller can be designed using fuzzy logic, with switching between different control laws depending on the variable values. Because of the complex spacecraft dynamics, there could be many ways to choose the switching conditions. However, the study of these options would be expensive and time consuming. Instead, a two-window nonlinear controller was designed which only changes the control law on the basis of the absolute value of the error in each channel. This was done by passing the errors via saturation-dead zone windows so that smooth transition between the control laws was provided. The resulting control law is nearly perfect for the despin function and as well for the cruise mode. The transient response for this controller is shown by the curve (2). The de-spin time was reduced by 20°/0.

The two-window controller performs better and is at the same time more robust than the original linear controller, with larger stability margins for large error mode when the cross-axis coupling is the largest.

This example shows that even for complicated plants with multi-channel coupled nonlinear feedback loops, a nonlinear two-window controller using only the error in individual channels for changing the control law provides nearly time-optimal performance, substantially better than that of linear controllers.

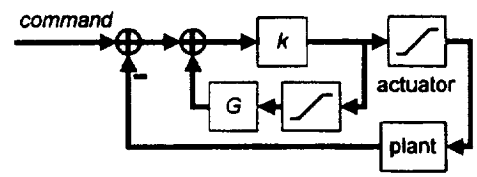


Fig. 9. Cassini main engine gimbal controller

Example 2. [n Cassini spacecraft, the main engine (thruster) is gimballed. The trajectory maneuvers can be performed by articulating the engine in x- and y-directions. The block diagram for the x-axis controller is shown in Fig. 9.

The error is processed by a two-window nonlinear controller which has a local feedback path comprising a saturation link and the guidance filter G. Small errors are multiplied by the transfer function $k/(1 - G)$. Large errors are multiplied by k. Intermediate amplitude errors are processed by the nonlinear compensator which is an intermediate between the two linear compensators. The two-window controller is of the kind described in [2,3]. It eliminates the wind-up, allows for large disturbance rejection, and assures asymptotic global stability.

The Cassini attitude control employs thrusters (without PWM). The plant is close to a pure double integrator, although there may be flex modes at high frequencies. The thrusters are not throttled and not modulated, the torque is positive or negative some fixed value, or zero (similar to a 3-position relay.) These controllers also employ two-window nonlinear compensators.

Example 3. Temperature controller for the mirrors of the Narrow View Camera of Cassini spacecraft. The camera is a small telescope shown in Fig. 10(a). The primary and the secondary mirrors of the camera must be kept at approximately same temperature in order for the mirror surfaces to match each other, and the image in the focal plane to be clear. Fig. 11 shows Bode diagrams for three [parallel channel compensator. The low-frequency (LF) channel is preceded by a saturation element which constitute amplitude window for the feedback error. The transient response to a step input is shown in Fig. 12. The time-response of the heater power shows that the controller is nearly time-optimal. The controller was described in more detail in [1].

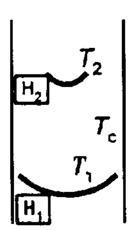


Fig. 10. Narrow Angle Camera

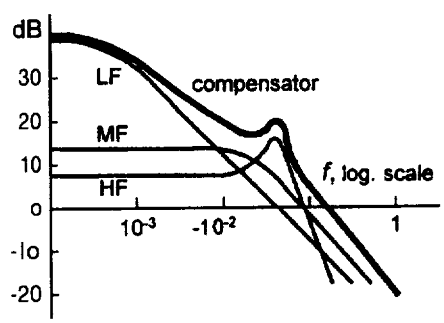


Fig. 11: Parallel-channel compensator responses

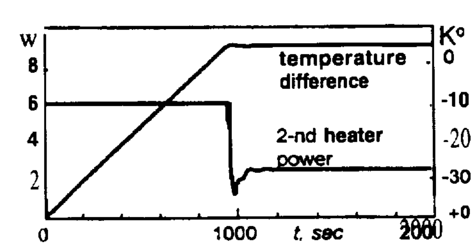


Fig. 12. Step response for thermal controller

Example 4. An tunnel-effect accelerometer [6] is shown in Fig. 13(a). The proof mass and the soft springs it is suspended on are etched of Silicon. The position of the proof mass is regulated by electrostatic forces between the proof mass and the upper and lower plates. The accelerometer uses tunnel effect sensor of the proof mass position. The voltage on the lower plate equals the voltage on the upper plate plus some bias. It can be shown that with proper bias voltage, the upper plate voltage is proportional to the measured acceleration.

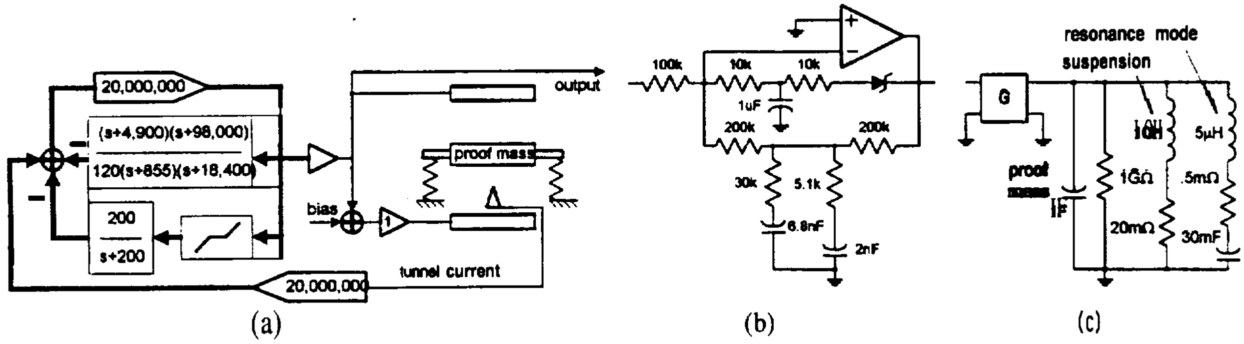


Fig. 13. Silicon accelerometer block diagram (a), compensator (b), and electrical analog plant model (c)

To achieve the desired accuracy, the feedback in the proof mass control loop must be larger than 100 dB at frequencies up to 5 Hz. The feedback crossover frequency f_b is limited by the dynamics (structural resonances) of the proof mass and suspension system to less than 3 kHz.

The tunnel current is the exponent of the inverse value of the tunnel effect gap. The normal value of the gap is approximately 6 Angstrom, but initially, the gap can be smaller, the tunnel current much larger, the derivative of it (the tunnel sensor sensitivity) also larger, and the loop gain bigger than nominal. The system global stability is provided by using a NDC with dead-zone in the local feedback path.

The mechanical plant might have some resonance modes with uncertain frequencies over 500 Hz. The quality factor of the resonances is not higher than 20, i.e. 26 dB.

The compensator is shown in Fig. 13(b). The dead-zone element was chosen non-symmetrical (a Zener diode) since the characteristic of the tunnel effect sensor is also non-symmetrical. For low level signals the Zener is not conducting, and the compensator response is determined by the lower feedback path. Two series RC circuits shunting the feedback path provide two leads giving sufficient phase stability margins over the range 200 to 3000 Hz. The Bode diagram and the Nyquist plot for signals of small amplitudes simulated in SPICE are shown in Fig. 13. When the signal is exceeding the Zener threshold, the diode opens and the upper feedback path, which is a low-pass, reduces the compensator gain at lower frequencies by approximately 30 dB. This gain reduction reduces the slope of the Bode diagram, substantially increases the phase stability margin at frequencies below 200 Hz, and improves the transient response of the closed loop which is important since the acquisition range of the tunnel effect sensor is very narrow, only about 15 Angstrom.

This controller provides global stability with the loop phase shift of π at frequencies where the loop gain is large, eliminates windup, reduces the overshoot, and increases the acquisition band of the tunneling condition. The tunnel effect is an exponential function, and if the feedback loop was initialized when the distance in the tunnel sensor gap was much smaller than normal, then, the loop gain is much larger, and the system would become unstable if it were not for the gain reduction by the NDC.

Example 5. A small parabolic antenna tracking the Earth is placed on Pathfinder Mars Lander. Two identical brushless motors with internal analog rate feedback loops articulate the antenna in two orthogonal directions. The motors are controlled by two independent identical SISO controllers.

The sampling frequency is 8 Hz. Were the delay caused only by the sampling, the crossover frequency would be $f_s/5 \approx$ 1.6 Hz. However, since the computer must handle not only the motor control loops but other higher priority tasks, there is an additional 500 msec delay caused by four real time interrupt (RTI) delays, 125 msec each. Also, due to limited bandwidth of the analog rate controllers for the motors (already designed), the motors have 50 msec delay. Since the total delay is not only 62.5 msec (of sampling) but 62.5 + 500 + 50 $\approx$ 600 msec, the realizable crossover frequency is lower in proportion to this delay, i.e. $f_b <$ 1.6.62.5/600 $\approx$ 0.17 Hz.

The controller includes two cascaded linear links, $C_1 + 1$ and C_2. A saturation link placed in front of C_1 makes the transfer function of the compensator dependent on the signal level. When the signal level is below the saturation threshold, the compensator transfer fiction is $(C_1 + 1)C_2$. When the signal is high, the compensator transfer function reduces to C_2.

For small signal amplitudes, the compensator function is $C = (C_1 + 1)\, C_2$ where C_1 is a single-pole low pass filter Cl = 2.5/(0.0833 +s) and C_2 is a lead link, $C_2 = (0.106 + s)/(2.23 + s)$. The asymptotic gain frequency responses of the compensator links are shown in Fig. 14. The digital compensator equations C_1ˉ(O. 15 + O. 15/z)/(1 - 0.99/z), C_2 = (0.9 - 0.8883 /z)/(1 - 0.75/2).

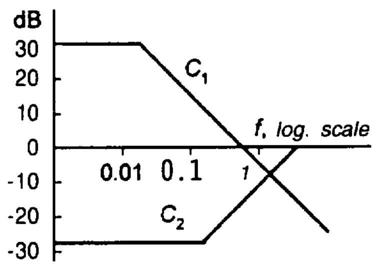


Fig. 14. Asymptotic Bode diagrams for compensators

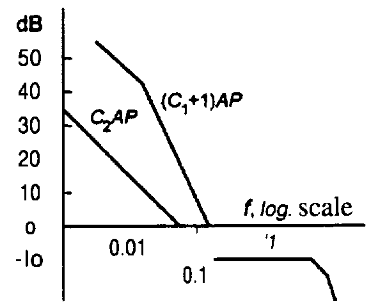


Fig. 15. Open loop asymptotic Bode diagrams of the compensators for small error (upper curve) and large error (lower curve)

The asymptotic loop gain frequency responses are shown in Fig. 15 for the case of both C_1 and C_2 operational, and for the case of C_1 = O (lower curve). The Bode step is very long because of the necessity to compensate for large time delay of up to 7 RTI, and to reduce or eliminate the overshoot,

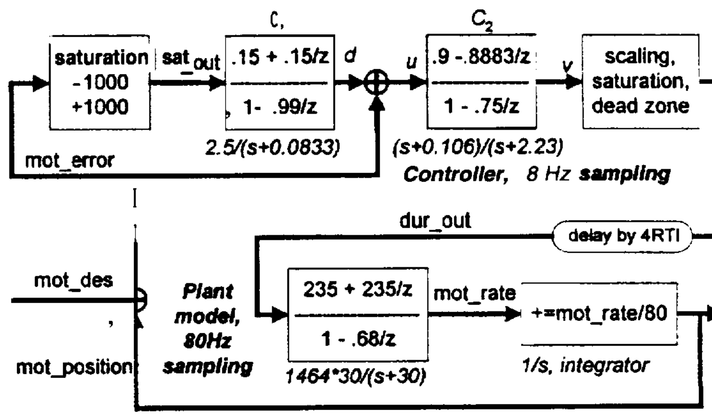


Fig. 16. Motor controller flow chart

The simplified feedback loop block diagram is shown in Fig. 16. The block diagram includes saturation in the higher gain, low frequency path; linear links C_1 and C_2; a scaling block that has saturation and a dead zone; the delay block; and the model of the plant (of the motor with its analog control electronics). The variable dur_out is the duration of the motor to be on during the sampling period of 125 msec. The motor is rate-stabilized by an analog loop with 30 msec rise-time. The motor transfer function is therefore that of an integrator (the angle of rotation is proportional of the time the motor is on) with an extra pole caused by the limited bandwidth of the analog rate loop.

Conclusion. Multiwindow controllers employed in space systems designed at JPL outperform conventional linear controllers and **simplify** the commanders.

We believe that such controllers should replace linear controllers in most of control systems, students must be taught that, generally, the controllers must be nonlinear, and newer control schemes should be compared to **multiwindow** controllers (instead of to linear controllers) to determine their advantages. (For example, it is shown in many examples that some **fuzzy** logic and **neural** network controllers perform substantially better than the best linear controllers, but comparisons to well designed **multiwindow** controllers are typically missing.)

Acknowledgment. This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The authors acknowledge contributions of Dr. P. Enright (who, in particular, designed the Cassini main engine articulation controller and the attitude thruster controller) and Drs. D. Bayard and M. Mesbahi for technical discussions.

References

1. P. J. Enright, F. Y. Hadaegh, B. J. Lurie, Nonlinear Multi-window Controllers. AIAA Guidance Navigation and Control Conference, July 29-31, 1996, San Diego, CA.
2. B. J. Lurie, P. J. Enright. *Classical **Feedback Control**,* manuscript, 1996.
3. B. J. Lurie, Absolutely Stable Feedback System with Dynamic Nonlinear Corrector. "Proc. IEEE", vol. 70, no. 8, 1982.
4. B. J. Lurie, Absolutely-Stable Nyquist-Stable Nonlinear Feedback System Design. "Intern. J. Control", vol. 40, no, 6, 1984.
5. B. J. Lurie, ***Feedback Maximization.*** Artech House, Dedham, MA: 1986.
6. B. P. Dolgin, F. T. Hartley, B. J. Lurie, P.M. Zavracky, Electrostatic Actuation of a Microgravity Accelerometer, 43rd National Symposium of American Vacuum Society, Microelectronical Mechanical Systems Topical Conference, Philadelphia, Pa 14-18 Oct. 1996.

URC97163

A Long Range Science Rover for Future Mars Missions

Samad Hayati*
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Dr., Pasadena, CA 91109
Samad.Hayati@jpl.nasa.gov, (818) 354-8273

ABSTRACT

This paper describes the design and implementation currently underway at the Jet Propulsion Laboratory of a long range science rover for future missions to Mars. The small rover prototype, called Rocky 7, is capable of long traverse. autonomous navigation. and science instrument control, carries three science instruments, and can be commanded from any computer platform and any location using the World Wide Web. In this paper we describe the mobility system, the sampling system, the sensor suite, navigation and control, onboard science instruments, and the ground command and control system.

1. INTRODUCTION

Even prior to the recent discovery of the possibility of past life on Mars by a research team of scientists at the Johnson Space Center and at Stanford University, NASA had planned six missions to Mars for 2001.2003, and 2005. Currently NASA is replanning its Mars exploration strategy to develop strategies that lead to one or more sample return missions. Since samples must be examined using onboard science instruments and collected from a variety of sites, rovers will play a crucial role in these missions. These rovers will traverse to sites separated by several kilometers and place instruments against outcrops or loose rocks, search an area for a sample of interest, and collect rocks and soil samples for return to Earth. Our research objectives are to develop technologies that enable such scenarios within the mission constraints of mass, power, volume, and cost.

Our goal is to develop technologies that overcome limitations detailed shove as well as to introduce new capabilities currently not supported. These are:

- Increase rover autonomy so that the number of science experiments per uplink command is increased, resulting in more science data. This involves increased autonomy for rover navigation to reach science targets, autonomous confirmation of reaching such targets, and use of sensory information to autonomously perform manipulation and science instrument placement and pointing.
- Develop the ability of the microrover to traverse long distances by integrating a celestial sensor (e.g., sun sensor) to determine rover's orientation; and by developing a deployable mast mounted camera system to send panoramic images of the surrounding area to the ground control personnel.
- Integrate representative science instruments on to the rover and develop intelligent data reduction techniques to maximize the useful science return.
- Develop onboard resource analysis and decision making capability so that maximum science is returned for the available resources.
- Develop a distributed Internet based rover interface so that scientists can provide science experiment requests and the general public can view return images immediately.
- Test and validate these technologies in realistic settings and with planetary scientist participation.

This paper provides an overview of our prototype rover called Rocky 7 and describes our near-t.erm goats. Section 2 gives a description of the mobility system, the sampling arm, sensors, the perception system, the navigation technique, and the science instruments. Section 3 describes a new operator interface development that allows a

---

*Samad Hayati is the task manager. Following individuals comprise the team that has designed and developed the Rocky 7 rover: Dr. Richard Volpe, Dr. Bob Balaram, Dr. Richard Welch, Dr. Paul Backes, Mr. Robert Ivlev, Mr. Tim Ohm, Mr. Richard Petras, Mr. Sieve Peters, and Mr. Greg Tharp.

rover to be commanded from any location using the World Wide Web. Conclusions are given in Section 4. References are provided in Section 5.

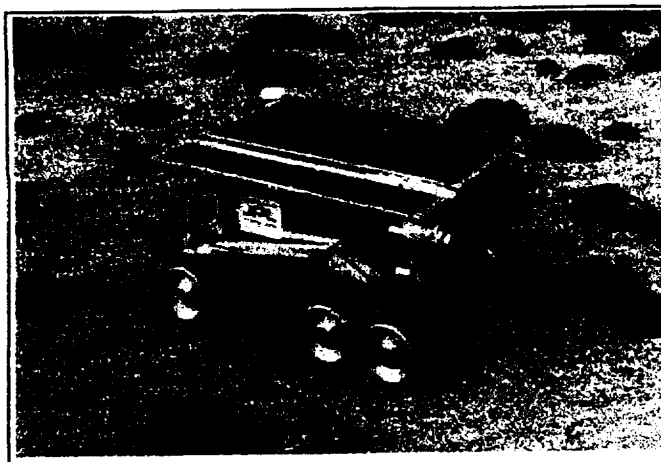


Figure 1. Rocky 7 rover in JPL Mars Yard shown with stowed mast and sampling arm

2. ROCKY 7 ROVER

In this section we provide the Rocky 7 rover configuration and detail the constituent components. Figure 1 shows Rocky 7 in the JPL Mars Yard. Mars Yard is a 15X25 meter outdoor test area that closely simulates Mars like terrain constructed based on statistical analysis of images taken by Viking Landers I and H.

One important consideration in developing Rocky 7 has been its flight relevance. This has severely constrained its size, mass, and power. The size of the rover is dictated by the size of the payload envisioned for future missions. Rocky 7 measures 48 cm wide, 64 cm long, and 32 cm high.

2.1 Mobility System

The mobility system is a modified Rocker-Bogie design used in previous rovers at JPL [2]. It consists of two rockers (hence the name "Rocky) hinged to the sides of the main body attached to six wheels two of which are steerable. This configuration requires two fewer actuators (total of eight) than previous version. Each rocker has a steerable wheel at one end and a smaller rocker at the other end. Two wheels are attached at the end of each of these small rockers. The main rockers are constrained in motion via a lever which is hinged at the end of the main body and its two ends are attached to the end of main rockers. This mechanism provides two important mobility characteristics for the rover. First, a wheel can be lifted verticality while other wheels remain in contact with the ground. This feature provides rock climbing capability to the rover. Rocky 7 can climb rocks 1.5 times its wheel diameter. Second, the vehicle can climb over rocks that span the width of the vehicle, using the smaller rocker and the two-wheel arrangement on the main rocker, even if a rock almost fits snugly between the front and middle wheels.

2.2 Sampling System

One significant improvement over previous Rocky series rovers is the incorporation of a sampling device on Rocky 7. The savings in actuators achieved by reducing the number of steerable wheels are used to develop the sampling system. This light weight (650 gm) sampling arm consists of a two- DOF manipulator (32 cm long) that is attached to the front of the rover and can reach 1 0 cm below the ground surface. When folded, it is in a horizontal position against the front of the rover. The arm has a two-DOF scoop mechanism and is designed to both dig and carry the samples. When scoops are rotated 180 degree backward the arm can grasp objects using the **back** side of the scoops. Figure 2 shows that sampling arm with acquired soil sample.

In addition to sampling function, the arm is used to deliver light to an optical fiber via a pair of mirrors. This is accomplished by configuring the scoops to a position and exposing a normally closed hole. The optical fiber carries the light (image) to a point spectrometer located inside the rover chassis,

Figure 2. Rocky 7 sampling arm and scoops

The arm is deployed for three different operations: digging, dumping, and spectrometer data acquisition. Before each deployment the rover checks for possible collision of the arm with the obstacles (rocks) using its on board stereo vision system and automatically positions itself to avoid them. For a dig operation, the vision system also processes the images of the area in front of the rover to determine if the ground is soil-like by analyzing the image texture and elevation information. It then deploys the arm and lowers it until contact is made with the surface by monitoring the arm motor current. After the dig operation, it positions the scoop that collected the sample and takes its image. It then compares this image against the one taken just before the dig operation. If it detects enough difference between these two images. the rover reports success and completes the dig operation by closing the scoops and stowing the arm. Otherwise it does an automatic dump, stows the arm, and reports failure. Similar autonomous checks are performed for a dump operation.

2.3 Sensors

Several sensors are used for navigation. A sun sensor developed by Lockheed Martin, called the Wide Angle Sun Sensor (WASS), provides heading information as a function of the rover's location and the time of day using an onboard real-time software module. In addition, an accelerometer is installed to provide pitch and roll information. The wheels are equipped with encoders for precise servo control and to estimate the rover s position. The rover is equipped with seven (extendible to eight) CCD cameras. Two at each end, for the perception system discussed in the next section, two at the end of a deployable mast and one in a close-up imager.

2.4 Perception System

To simplify the perception system hardware, Rocky 7 uses only a passive stereo vision [3] for hazard detection unlike its predecessor that used a laser striping system in conjunction with multiple monocular cameras to detect obstacles. The stereo vision system uses a pair of cameras with wide angle lenses to allow viewing of both the manipulator and its actions as well as to permit imaging of rocks and other hazards extending from near the rover to a little above the horizontal. Rocky 7 is equipped with six cameras that can be used for navigation: two in front of the rover, two in the back, and two on the mast.

One advantage of the stereo vision system is that it is easy to extend the system capability by adding additional cameras to the back side of the rover and use the existing infrastructure (i.e., frame grabbers and software) to perform collision avoidance.

2.5 Navigation System

Rocky 7 navigation strategy is based on operator waypoint designation and autonomous behavior based navigation to move to the specified targets [4, 5]. Operation starts with a command issued to the rover to take a panoramic image of the scene by obtaining several overlapping images. These images are then processed by a stereo vision software to obtain terrain maps on the ground. An interactive software allows one to select specific points (locations) on one of these images using a mouse. The software returns the position of this location as calculated by the stereo vision system and displays the coordinates of the point. If no valid coordinate exists for the particular point, the software indicates this to the operator. The operator continues this operation and builds a path which deems to be safe for the rover to traverse through in moving from its initial position to the target location.

Before each move, the rover takes a set of images and process them onboard and determines if there are obstacles that it must avoid. If there is no obstacle it moves a short distance and then stops and repeats the same operation.

If it determines that there is an obstacle. then it turns away by a fixed amount to the right or to the left depending where the obstacle is.

2.6 **Science Instruments**

An important objective of our research in developing rovers is to understand not only the mobility, navigation. and control issues, but to also consider problems associated with the integration of science instruments, their on board operation and data reduction. Currently Rocky 7 has three science instruments: a point reflectance spectrometer. a wide field of view spectral imager. and a close -up spectral imager.

The point reflectance spectrometer is on-board the rover chassis and its fiber optic path is integrated into the rover manipulator. This allows the spectrometer to be pointed at rock/soil targets from many different angles. Also included on the manipulator is a calibration target for taking reference data for the current illumination. In the near future a laser will be added to the fiber optic path so that the point of the spectra data can be illuminated and imaged to confirm exactly where the spectral data was taken,

The wide field of view spectral imager is developed by adding motorized filter wheels to the mast cameras. This filter wheel system is used to gather broad baud spectral data enabling color images to be constructed. The mast is a three degrees of freedom torso/shoulder/elbow articulated robotic arm enabling the cameras on the end to be positioned 1.4 meters above the ground as well as being able to pan and tilt the cameras to get the desired imagery. The cameras are shown in Figure 3.

The third instrument is a close-up imager that uses a monochrome camera and active lighting source. This is packaged as a 500g "dummy"' instrument representing an APX or Moessbauer spectrometer which would have to be placed against a designated target. The instrument is mounted at the end of the mast and the mast degrees of freedom are utilized to position instrument against rocks in front of the rover. Passive compliance is used to allow the instrument to orient itself normal to the target surface and contact sensors used to confirm placement.

2.7 Long Traverse

Rocky 7's mast and the sun sensor allow the rover to traverse long distances. The scenario for the operation of the rover consists of traversing in the indicated direction, using the sun sensor, and **periodically (e.g., -** 100m to 200 m) transmitting panoramic images to the ground station. These panoramic images are obtained by a pair of cameras mounted on a stowable mast that is carried by the rover. The ground station provides new commands to either to continue to traverse in the same direction or to change direction. If the site is of interest to scientists, site survey commands will be issued. For each site survey. a panoramic image is used to designate science targets and to specify science experiment parameters (such as angle of pointing a science instrument relative to a target, instruments distance from a target, and duration of data acquisition, etc.). The rover then autonomously performs the requested science experiment. Success of each science experiment is confirmed by the rover autonomously via executing specific tests for that experiment.

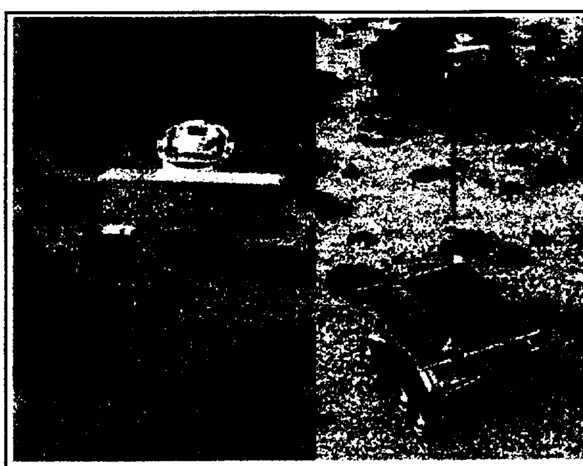


Figure 3. Rocky7 mast shown in the upright position (right) Spectral and Close-up imagers (left)

3.0 Advanced Operator Interface

We have developed a ground control station to remotely command the rover and receive data from it. The operational scenario is based on the rover down-linking science data and stereo panoramic image pairs. This data along with camera parameter information is used to develop terrain maps.

The interface is based on a World Wide Web (WWW) which consists of viewing an image taken by a rover camera. Through a mapping between this image and an elevation map discussed earlier an operator can point and click on any point on the image and obtain the coordinates of the point. This technique has been used before at JPL for target selection and waypoint designation successfully. This Web based version of the allows a scientists to select science targets in his or her home institution using any computer platform. He/she is also be able to describe the nature of a particular science

experiment to be performed at that point (pointing requirements, time required for data collection, data compression, etc.), This information is then sent electronically to a central station at JPL for consolidation and verification for flight rules for next day's mission and for uplinking to the rover. Figures 4 and 5 show the interface for remote target and waypoint selection.

Figure 4. WEB based interface. Right image shows waypoints selected. The left image hews a top view of the elevation map generated from panoramic images. The right image corresponds to one of the wedges shown in the left image. This interface can be accessed rom: ***http://robotics.jpl.nasa.gov/tasks/scirover/operator/wits/index.html***

Plans are also underway to provide panoramic elevation maps to clearly show the camera image in the context of the panoramic elevation map, The operator control station will also be able to show these in the context of descent imagery which is very important for scientists planning their global exploration strategy.

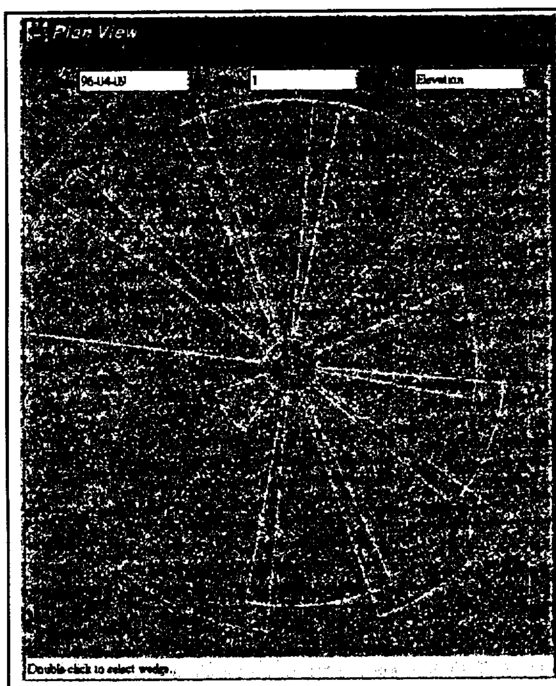


Figure 5. This image is the same image as shown in Figure 5 without the elevation map. The operator has an option of looking at the left image of Figure 5 with or without the elevation map.

In the future. we will perform feature segmentation and provide feature maps to identify landmarks for rover localization autonomously.

4. CONCLUSIONS

This paper has provided an overview of research on future Mars rovers covering navigation, perception, science instrument pointing and placement, and operator interface issues.

Although this research program covers many essential elements of Mars rovers, research related to materials, space qualified computers, communication hardware, thermal insulation. advanced mobility systems, and structures are being address by other tasks at JPL [6].

ACKNOWLEDGMENTS

Research described in this paper was carried out by the Jet Propulsion Laboratory, California institute of Technology, under contract with the National Aeronautics and Space Administration.

5. REFERENCES

[1] Matijevic J. *Mars Pathfinder Microrover - Implementing a Low Cost Planetary Mission Experiment.* Proceeding of the Second IAA International Conference on Low-Cost Planetary Missions, John Hopkins Applied Physics Laboratory, Maryland, USA, April 16-19, 1996, paper # IAA-L-0510

[2] Bickler D. *A New Family of JPL Planetary Surface Vehicles.* Missions, Technologies, and Design of Planetary Mobile Vehicles, pp. 301-306, Toulouse, France, September 28-30, 1992.

[3] Matthies L and Grandjean P. *Stochastic Performance Modeling and Evaluation of Obstacle Delectability with Imaging Range Sensors."* IEEE Transactions on Robotics and Automation, Vol. 10, pp. 783-791, December 1994.

[4] Volpe R, et al. *The Rocky 7 Mars Rover Prototype.* Workshop on Rovers, IEEE Robotics and Automation Conference, Minneapolis, Minnesota, April 22-28, 1996.

[5] Gat E, et al., *Behavior Control for Robotic Exploration of Planetary Surfaces.* IEEE Transactions on Robotics and Automation, Vol. 10, pps. 490-503, 1994.

[6] Weisbin C. *Rovers and Telerobotics Technology* Program. Robotics and Mars Exploration Program Office Publication. JPL.

Robust Flight Path Determination for Mars Precision Landing Using Genetic Algorithms

David S. Bayard and Hamid Kohen

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109-8099

ABSTRACT

This paper documents the application of genetic algorithms (GAs) to the problem of robust flight path determination for Mars precision landing. The robust flight path problem is defined here as the determination of the flight path which delivers a low-lift open-loop controlled vehicle to its desired final landing location while minimizing the effect of perturbations due to uncertainty in the atmospheric model and entry conditions. The genetic algorithm was capable of finding solutions which reduced the landing error from 111 km RMS radial (open-loop optimal) to 43 km RMS radial (optimized with respect to perturbations) using 200 hours of computation on an Ultra-SPARC workstation. Further reduction in the landing error is possible by going to closed-loop control which can utilize the GA optimized paths as nominal trajectories for linearization.

1. INTRODUCTION

In this study, GAs are applied to optimizing a nonlinear simulation of descent dynamics of a low-lift vehicle during planetary (i.e., Mars) entry. The basic idea is to find a flight path which comes closest to a desired landing position, yet is robust to expected perturbations in the trajectory. Such a ***robust*** flight path is found by minimizing a quadratic cost function representing the landing miss distance, over several realistically perturbed trajectories. The most important perturbations are the error in the initial entry conditions, and uncertainties in the atmospheric density. In order to vary the flight path, the initial flight path angle is chosen as a free parameter, and the vehicle angle-of-attack is controlled as a function of time. The control of the angle-of-attack is accomplished using the center-of-mass (COM) relocation concept put forth by D. Boussalis of JPL [1]. The COM relocation concept is important because it allows considerable control authority during the atmospheric entry phase to minimize landing errors, yet it is applicable to low-lift Mars Pathfinder type aeroshells (i.e., with lift-to-drag ratio L/D = 0.3). This avoids the need for designing higher lift (and much more expensive) vehicles. For simplicity the entry dynamics have been restricted to planar motion, and the landing error is defined at 10 km altitude where the parachute opens rather than at ground level. This paper is an abridged version of a longer report [11].

2. CONTROL ACTUATION

The control actuation scheme will be based on center-of-mass (COM) relocation, as outlined in Boussalis [1]. In this approach, a proof-mass is moved inside the vehicle so that the COM is relocated as a known function of time. The COM relocation acts to shift the dynamic equilibrium of the vehicle such that the angle-of-attack is changed. In particular, the equilibrium angle-of-attack value varies as an explicit known function of the COM relocation. Hence, even though one is moving a proof-mass, the control can be thought of as commanding a desired angle-of-attack. Since the angle-of-attack acts to change the amount of lift or drag on the vehicle, it provides a means to effect the propagation of the flight path.

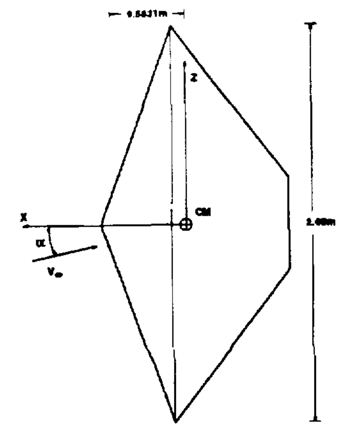


Figure 1 Low lift Mars Pathfinder type aeroshell

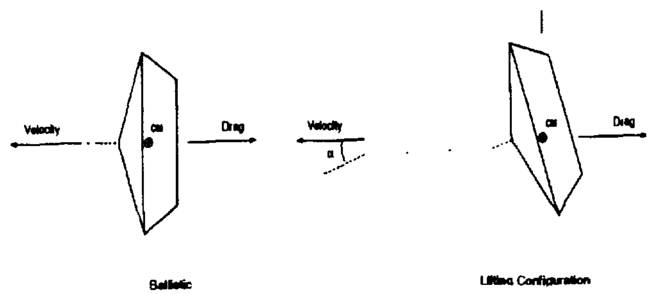


Figure 2 Center-of-mass relocation scheme to control lift vector

3. ROBUST FLIGHT PATH PLANNING MODEL

For the purposes of this study, the "landing error" is defined as the RMS error in the desired terminal ground track location over a collection of 5 simulated paths, i.e.,

$$J = \sqrt{\sum_{i=A,B,C,D,E} (S_{xd}-S_{xi})^2 + (S_{yd}-S_{yi})^2} \qquad (1)$$

where S_{xd}, S_{yd} (specified later) are the desired ground track at the terminal time, and S_{xi}, S_{yi} are the actual ground track at the terminal time.

For the purpose of evaluating the RMS error ***J, the*** 5 simulations (A, B, C, D, and E) are performed per control profile to determine the effect of perturbations on the flight path. Parameter perturbations associated with A, B, C, D and E are shown in Table 1 and Figure 3. These perturbations reflect the major sources of error in the descent phase which are due to ***uncertainty*** in the atmospheric parameter beta, and uncertainty in delivery to the specified initial flight path angle gamma(0) (i.e., the entry corridor).

Three scenarios are addressed for optimization of the flight path:

Scenario 1: Two Point Boundary Value Problem, Constant Control
Find the control (i.e., the entry condition gamma0, and fixed COM offset dz) that under perfect knowledge and no disturbances, places the vehicle at the desired final position (in terms of its desired ground track) at the terminal time (i.e., the time instant at which the altitude is 10 km, and the parachute deploys). Apply this control to the 5 perturbed trajectories to calculate RMS landing error ***J***.

Scenario 2: Robust Flight Path Determination, Constant Control
Find the control (i.e., the entry condition gamma0, and fixed COM offset dz) that optimizes the RMS landing error ***J*** at the terminal time over the 5 perturbed trajectories.

Scenario 3: Robust Flight Path Determination, 5th Order Control
Find the control (i.e., the entry condition gamma0, and the COM offset dz as a 5th order Chebchev polynomial function of time dz=u(t)=Trun[a0+a1*c1(t)+...a5*c5(t)]), that optimizes the RMS landing error *J* at the terminal time over the 5 perturbed trajectories. Acontrol contraint on dz to +/-.08 m is enforced by the operator Trunc[], which truncates the Chebychev polynomial when it exceeds these thresholds.

Note that by minimizing the RMS landing error *J*, one is not only delivering the vehicle to its desired final position under nominal conditions, but is also minimizing the effect of perturbations on the actual flight path. This is the essence of the robust flight path planning problem.

Table I Perturbed Parameters for Simulation

Indv Runs	beta	gamma(o)
A	1.00*beta0	gamma0 + 0.0
B	1.25*beta0	gamma0 + 0.2
c	0.75*beta0	gamma0 + 0.2
D	0.75*beta0	gamma0 -0.2
E	1.25*beta0	gamma0 -0.2

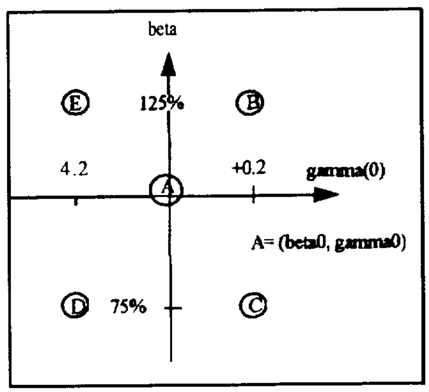


Figure 3 Flight path angle (gamma0) and atmospheric (beta) perturbations

The kinematics and dynamics of the vehicle. during descent are described by the a system of differential equation which can be found in [1][11].

4. GENETIC ALGORITHM IMPLEMENTATION

The Genetic Algorithm Toolbox [7] is used to solve the three scenarios posed in the previous section. For this purpose, the chromosomes are set up as shown in Table 2 and the initial conditions are given in Table 4. The desired final landing location is specified as, $S_{xd} = 556.1$ km and $S_{yd} = 976.65$ *km*.

Table 2 Chromosome Coding

Chromosome	Range value	Precision
gamma0 (degree)	-9 to -17	15 bit
dz (m)	-0.08 to 0.08	15 bit
ai, i=0,...,5	-0.08 to 0.08	15 bits

Table 3 Summary of Computational Requirements

Scenario	# Individuals per population	# Generations	Machine	Memory RAM	Speed	Hours
I	10	20	Pentium	16 Meg	133 Mhz	172
II	20	27	Ultra SPARC	132 Meg	143 Mhz	90
III	20	60	Ultra SPARC	132 Meg	143 Mhz	200

Table 4 Initial States (all scenarios)		
Altitude	125.0	kilometer
Longitude theta	0.0	degree
Latitude phi	-10.0	degree
Velocity	7.5	kilometer/sec
Flight path angle, gamma0	Evolved	degree
Azimuth (heading) angle, psi	60.0	degree
Pitch rate, q	0.0	degree/sec
Pitch	gamma0+alpha0	degree
alpha0	$-C_{m0z} * dz(0) / C_{m\alpha}$	degree
Sx Ground track	0.0	kilometer
Sy Ground track	0,0	kilometer

5. ANALYSIS OF THE RESULTS

The results of all three scenarios are tabulated in Table 5.

Table 5 Summary of Results

	gamma0 (degree)	dz (cm)	Landing Error - RMS Radial (km)
Scenario I	Evolved -12.54	Evolved -0.03713	111.68
Scenario II	Evolved -13.58	Evolved -0.0610	75.825
Scenario III	Evolved -12.5080	Chebychev a0 = 0.0145 al = 0.04096 a2 = -0.0690 a3 = 0.0260 a4 = 0.0530	43.3855

For comparison purposes, the landing error plots for Scenarios I,II and III are organized from left to right in *Figure 4*. As expected the RMS landing errors decrease from left to right with increasing control authority.

a. Scenario I b. Scenario 11 c. Scenario 111

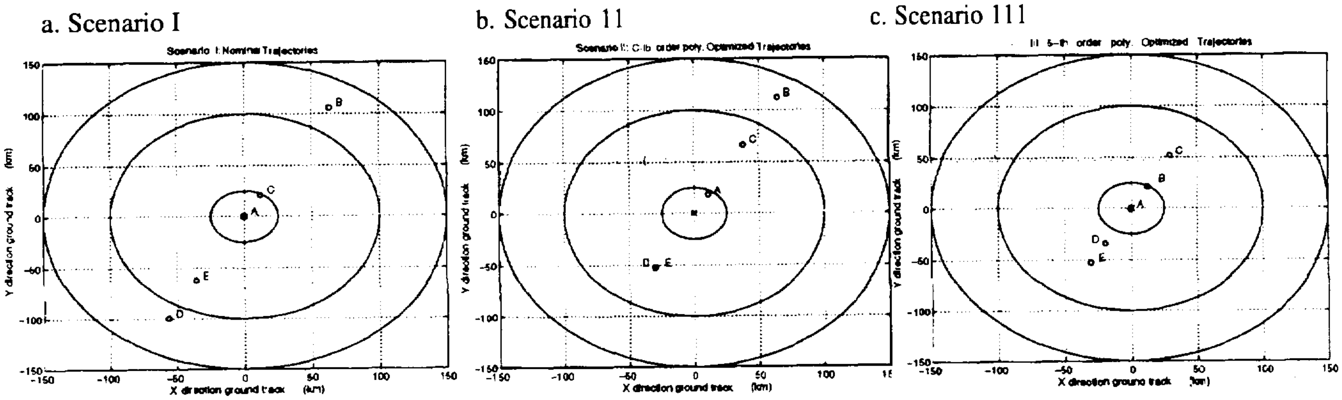


Figure 4 Summary of landing errors for all scenarios

The improvement in going from Scenario I (111 km) to Scenario II (76 km) is to be expected since Scenario I was not optimized with respect to the perturbed trajectories while Scenario II was. The improvement in going from

Scenario II (76 km) to Scenario III (43 km) is also expected since Scenario III is a generalization of Scenario II in terms of progressing from a zeroth order polynomial to a 5th order polynomial control representation.

a. Scenario I b. Scenario II c. Scenario III

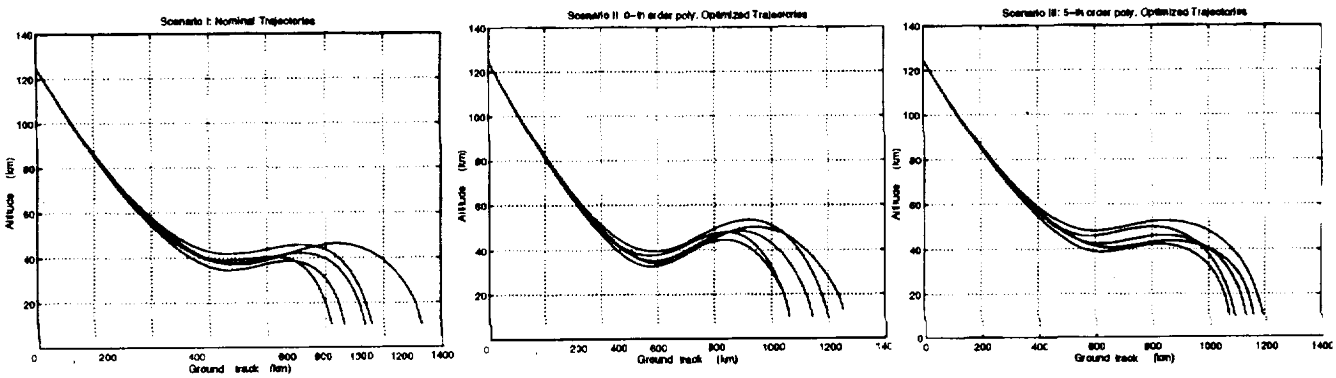


Figure 5 Summary of altitude paths for all scenarios

It is instructive to compare the altitude plots of the three Scenarios in Figure 5. It is seen in Scenario III how the GA successfully reduces landing error by making the perturbed flight paths coalesce.

The flight path determined by GA for the 43 km (Scenario III) result is very interesting and suggests a new "bounce and plop" strategy for precision landing. In order to study this strategy in more detail, the altitude and control signal dz=u(t) for Scenario III are plotted on the same x-axis (i.e., versus time) in Figure 6. The scale for the control signal has been converted to mm to allow sharing of the same y-axis. It is seen that the "bounce" is induced by lowering the COM (i.e., dz=u(t)) to its maximum negative location of u= -.08 m (i.e., maximum positive lift), at approximately 10 seconds. Note that the bounce does not take effect until the atmosphere is sufficiently dense at an altitude of 40 km (occurring at approximately 75 seconds), to create a significant lift effect. The "plop" is induced by raising the COM location to its maximum positive location of u= +.08 m (i.e., maximum negative lift), at approximately 135 seconds. Again, the negative lift is seen to take effect when the atmosphere becomes sufficiently dense at an altitude of 40 km (occurring at approximately 200 seconds). This overall approach forces the perturbed trajectories to coalesce, which effectively reduces landing error.

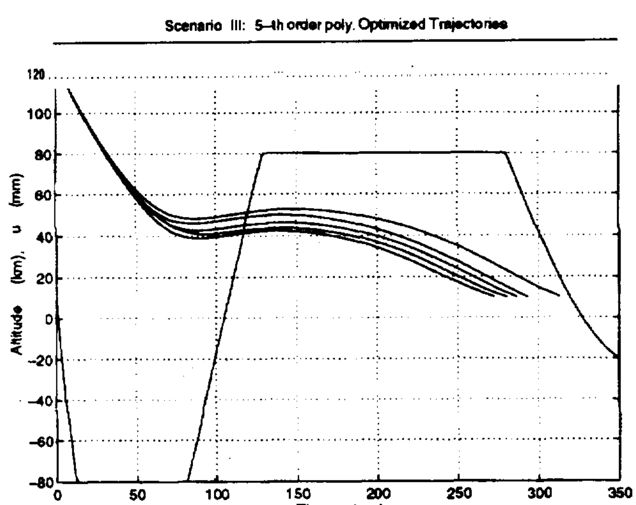


Figure 6 Superposition of vehicle altitude and control signal dz=u(t)

6. CONCLUSIONS

A genetic algorithm was applied to the problem of robust flight path determination for Mars precision landing. The notion of a robust flight path appears to be new, although it is a natural statement of what is desired in many open-loop control scenarios. In this study, the objective of the robust flight path problem was to determine the flight path which delivers a low-lift open-loop controlled vehicle to its desired final landing location while minimizing the effect of certain realistic perturbations.

The results of the study can be summarized as follows. When the control (i.e., the COM location) is chosen constant with time and the flight path is optimized with respect to the nominal trajectory, the resulting landing error is 111 km RMS radial. When the control is chosen constant with time and the flight path is optimized over perturbed trajectories, the landing error is reduced to 76 km RMS radial. When the control is allowed to vary as a fifth order polynomial and the flight path is optimized over perturbed trajectories, the landing error is 43 km. The trajectory determined by GA for the 43 km result is very interesting and suggests a new "bounce and plop" strategy for landing.

The major computational bottleneck for this study was in evaluating the objective function (or equivalently, the "fitness") for each individual in the population, since it required integrating the kinematics and dynamics of motion. For implementation purposes, it was necessary to trim down the GA implementation to a reduced population of 20 individuals and no more than 60 generations, requiring approximately, 20*10*60/60=200 hours of computation on an Ultra SPARC computer. Methods to reduce the computation time would be greatly beneficial.

Results indicate that even though genetic algorithms may require long processing times, they are fairly easy to program, and can provided useful solutions to complex optimization problems, such as those associated with problems of robust flight path planning, and spacecraft autonomy.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. D. Boussalis of JPL's Automation and Control section for technical advice and support. The research described in this document was carried out by the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration. The research was performed for the Microspacecraft Systems Technology Office of the JPL Technology and Applications Programs Directorate by personnel from the JPL Engineering and Science Directorate, and it was sponsored by the Spacecraft Systems Division of the NASA Office of Space Access and Technology.

REFERENCES

[1] D. Boussalis, ***Investigation of the Longitudinal Motion of Low-Lift Vehicles,*** JPL Internal Document, Engineering Memorandum EM 3456-96-002, May 7, 1996

[2] J. Koza, D.E. Goldberg, D.B. Fogel and R.L. Riolo (Eds.), ***Genetic Programming 1996,*** Proceedings of the First Annual Conference, Stanford University, July 28-31, 1996

[3] J. Holland, ***Adaptation in Natural and Artificial Systems, The*** University of Michigan Press, Ann Arbor, 1975.

[4] D. Farless, "Mars Precision Landing Study Team Summary Report for FY95," JPL Internal Document, Interoffice Memorandum IOM 3 12/96.6-002, January 23, 1996.

[5] M. Srinivas and L.M. Patnaik, ***Genetic Algorithms: A Survey,*** IEEE Computer Magazine, Vol. 27, No. 6, pp. 17-27, June 1994.

[6] J.L.R. Filho, P.C. Treleaven and C. Alippi, ***Genetic-Algorithm Programming Environments,*** IEEE Computer Magazine, Vol. 27, No. 6, pp. 28-45, June 1994

[7] A. Chipperfield, P. Fleming, H. Pohlheim, C. Fonseca, ***Genetic Algorithm Toolbox, User's Guide,*** Version 1.2, Dept. Automatic Control and Systems Engineering, University of Sheffield.

[8] D. E. Goldberg, ***Genetic Algorithms in Search, Optimization and Machine Learning,*** Addison Wesley Publishing Company, January 1989.

[9] J. E. Baker, ***Reducing bias and inefficiency in the selection algorithm,*** Proc. ICGA 2, pp. 14-21,1987.

[10] Z. Michalewicz, Genetic Algorithms+ Data Structures= Evolution Programs. AI Series. Springer-Verlag, New York, 1994.

[11] D.S. Bayard and H. Kohen, ***Genetic Algorithms for Spacecraft Autonomy: Flight Path Optimization for Mars Precision Landing.*** JPL Internal Document JPL D-13900, Volume 6, October 11, 1996.

URC97165

RECONFIGURABLE POINTING CONTROL FOR HIGH RESOLUTION SPACE SPECTROSCOPY

5/64/39

David S, Bayard and Tooraj Kia
Jet Propulsion Laboratory
California institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109

Jeffrey Van Cleve
Department of Astronomy
Cornell University
Ithaca, NY 14853

Abstract

In this paper, a pointing control performance criteria is established to support high resolution space spectroscopy. Results indicate that these pointing requirements are very stringent, and would typically be difficult to meet using standard 3-axis spacecraft control. To resolve this difficulty, it is shown that performance can be significantly improved using a reconfigurable control architecture that switches among a small bank of detuned Kalman filters. The effectiveness of the control reconfiguration approach is demonstrated by example on the Space Infra-Red Telescope Facility (SIRTF) pointing system, in support of the Infrared Spectrograph (IRS) payload.

1 Introduction

Spectroscopy measurements are important for many types of scientific observations, and as a result are used in a wide variety of spacecraft payloads. For example, NASA's Space Infra-Red Telescope Facility (SIRTF), is expected to carry the InfraRed Spectrograph (IRS) payload to obtain various high-resolution spectrographs of interstellar, matter, planetary nebula and galactic nuclei.

High resolution spectroscopy depends on the accurate determination of the ratios of measured spectral lines. This requires that the flux obtained during measurement is not significantly degraded (i.e., offset) by motion of the image spot in the entrance slit during the exposure. Because of properties of the slit geometry and the imaging optics, the flux offset varies as a complicated function of both the pointing hiss and jitter [4][5][6].

In Section 2, a pointing control performance criteria is established to support high resolution space spectroscopy. In contrast to the case of imaging instruments which degrade (i.e., blur) primarily as a simple function of the jitter, the flux offset is shown to vary as a nontrivial function of both the pointing hiss and jitter [1] [4] [5] [6]. Due to this dependence on both bias and jitter, it is shown that typical pointing requirements needed to support high-resolution spectroscopy are quite stringent, and would typically be difficult to meet using standard 3-axis spacecraft control.

To resolve this difficulty, it is shown in Section 3 that performance can be significantly improved using a reconfigurable control architecture that switches among a small bank of detuned Kalman filters. The effectiveness of the control reconfiguration approach is demonstrated by example on the SIRTF pointing system, in support of the IRS payload. Conclusions are postponed until Section 4.

2 Pointing Requirements

2.1 Signal Diagram

A detailed signal diagram representing the spectroscopy requirements is shown in Figure 1. The quantity $w_o(\tau)$ represents the pointing process, which is assumed to be a second-order stationary Gaussian random process with mean w_b and variance σ_a^2. In painting control language, w_b is defined as the *bias* and σ_a is the long-term jitter, i.e., the RMS jitter associated with windows of infinite duration.

The pointing process $w_o(\tau)$ is expressed in units of arseconds, and is defined with respect to the slit center, For example, if w_o = O the image spot will be directly at the slit center, The coefficient A_2 of the square law has units of $(arcsec)^{-2}$ so that the quantities ξ and x are dimensionless. The coefficient A_2 *is* determined

by fitting curves depicting fractional flux offset versus position, and is in general a function of slit geometry, wavelength, and the optical design [4]. Only motion along the slit width (i.e., the dispersion direction) is considered in the analysis since performance is relatively insensitive to motion along the slit length. As a" result, all expressions will be in terms of single axis requirements resolved along the dispersion direction.

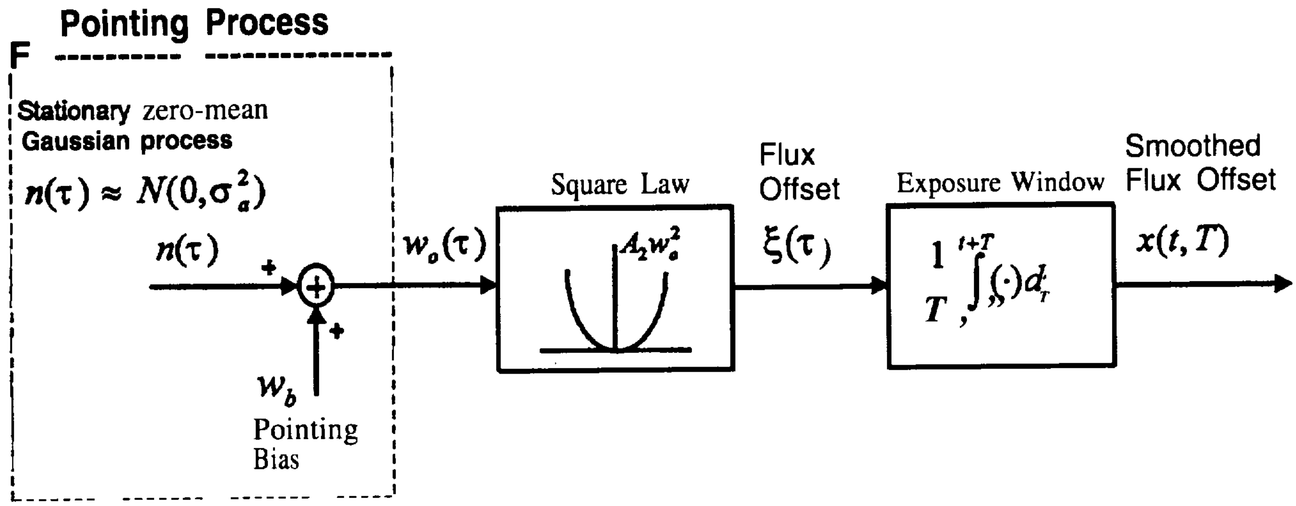


Figure 1: Signal diagram for spectroscopy pointing requirements

2.2 Statistical Analysis

The pointing control objective is to keep the image spot in the center of the slit by keeping the smoothed flux offset z small. Specifically, for accurate measured line ratios, it is desired that the probability of x exceeding a specified threshold d be less than a specified probability cr. Equivalently,

$$P_x(x \geq d) \leq \alpha \tag{1}$$

where $P_x(.)$ *is* the probability distribution of $x(t,T)$ in Figure 1. Since $x(t,T)$ is a stationary process in time, the probability P_x will not depend on t, but will in general be a function of the exposure time 'T.

Let $x_{1-\alpha}$ be the $(1-\alpha)$% percentile of the random variable x defined as follows,

$$P_x(x \geq x_{1-\alpha}) = \alpha \tag{2}$$

Then the pointing condition (1) can be equivalently written as,

$$x_{1-\alpha} \leq d \tag{3}$$

For infinite-time exposures (i.e., $T \rightarrow \infty$), the percentile $x_{1-\alpha}$ can be evaluated analytically as [1],

$$x_{1-\alpha} = A_2(\sigma_a^2 + w_b^2) \quad \text{valid for } T \rightarrow \infty \tag{4}$$

Unfortunately, for exposures of finite duration T, expression (4) is not valid, and the percentile $x_{1-\alpha}$ is much more difficult to evaluate. Hence, it will be replaced by an overbound $\bar{x}_{1-\alpha}$, which can be used to enforce (3) as follows,

$$x_{1-\alpha} \leq \bar{x}_{1-\alpha} \leq d \tag{5}$$

In [1], using Bienayme's inequality (Papoulis [2] pp. 115), such an overbound is obtained of the form,

$$\bar{x}_{1-\alpha} = \frac{A_2}{\sqrt{\alpha}} \cdot [3(\sigma_a^2 + w_b^2)^2 - 2w_b^4]^{\frac{1}{2}} \quad \text{valid for any } T \geq 0 \tag{6}$$

Using (6) in (5) and rearranging gives the pointing requirement,

$$[3(\sigma_a^2 + w_b^2)^2 - 2w_b^4]^{\frac{1}{2}} \leq \sqrt{\alpha} \cdot \frac{d}{A_2} \tag{7}$$

It is emphasized that (7) is very different from requirements for imaging instruments which avoid smearing by constraining the allowable RMS jitter over a window of specified duration (cf., [3]). In contrast, requirement (7) simultaneously constrains both the pointing hiss and jitter.

2.3 Three-Axis Control

As a realistic example, consider the values $\alpha = .05$ (for 95% confidence), $A_2 = .13$, and d = .07 (i.e., from the SIRTF IRS [4]). Substituting these values into (7), assuming equal contributions from bias and jitter gives,

$$\sigma_a = |w_b| \leq .195 \text{ arcseconds} \tag{8}$$

While it may be possible to meet the jitter requirement by taking advantage of optimal filtering and a good gyro/tracker combination, these requirements are quite stringent from the bias point of view. For example, a bias error of .2 arcseconds is by itself smaller than the accuracy of most available star trackers, and in addition there will be many other factors contributing to the overall pointing bias.

3 Reconfigurable Control

3.1 Architecture

It was seen that pointing requirements for high resolution spectroscopy are difficult to meet using standard 3-axis spacecraft control. An alternative approach based on a reconfigurable controller is proposed in this section, The basic idea is to place the image spot into the slit using a precision incremental maneuver on gyros, which avoids the bias error associated with the star tracker.

The proposed reconfigurable control architecture is shown in Figure 2. Here, KF1 and KF2 are Kalman filters which have been detuned to have time constants τ_1 and τ_2, respectively. KF1 and KF2 are both driven by the measured position quaternion q_m and measured 3-axis rate ω_m, while KFg is the optimal Kalman filter designed only with a rate measurement input. In this scheme, KF1 and KF2 are free running filters, while KFg is initialized by KF1 at time $t = 0$.

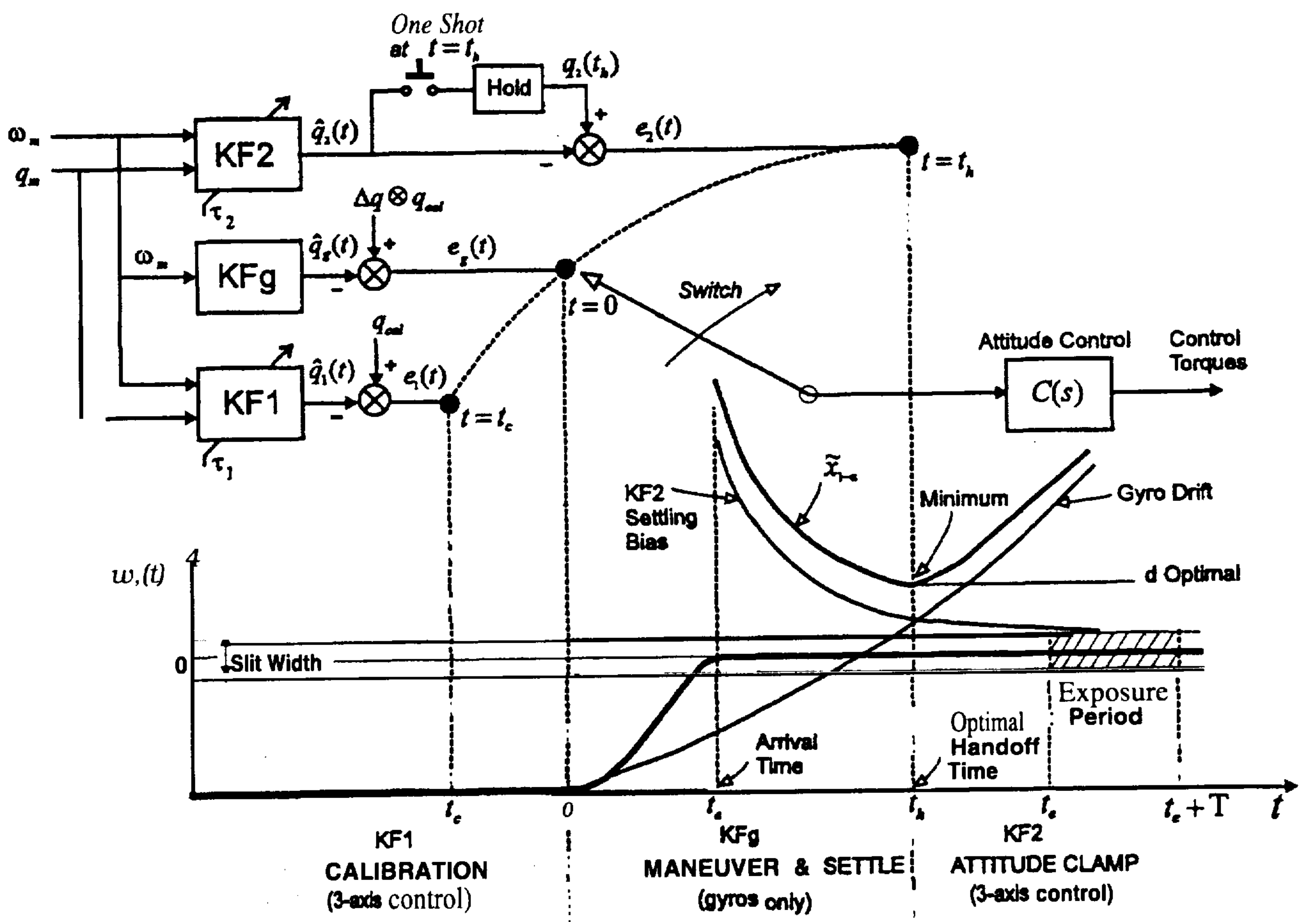


Figure 2: Reconfigurable control architecture for high-resolution spectroscopy

3.2 Handoff Description

As shown in Figure 2, the error signal which drives the attitude controller is taken from KF1 at time $t = t_c$, and is switched to KFg at time $t = O$, and is switched to KF2 at time $t = t_h$. It is assumed that the telescope and star tracker are in different frames, and that the body frame is the star tracker frame. Details of the particular handoff sequence are given below,

1. Point telescope to calibration source at attitude q_{cal} by nulling control error $e_1(t)$ associated with KF1.
2. Calibrate frame misalignment between tracker and telescope using calibration source (as imaged on a detector in the telescope frame) during time interval $t_c \leq t \leq 0$ while holding attitude on KF1.
3. Calculate incremental offset Δq in body frame needed to put a target source with known J2000 coordinates into center of spectroscopy slit.
4. At time $t = O$, command the attitude $\Delta q \otimes q_{cal}$ [where $\otimes$ denotes standard quaternion multiplication), and null control error $e_g(t)$ associated with KFg to implement maneuver.
5. Target arrives at slit at time $t = t_a > 0$.
6. At $t = t_h \geq t_a$ sample the "one-shot" to clamp attitude estimate associated with KF2, and null the control error $e_2(t)$.
7. Hold attitude by nulling $e_2(t)$ until spectroscopy exposure of duration T is completed.

It is emphasized that the attitude estimate from KF2 is clamped at time t_h to generate the control error $e_2(t)$ to be nulled. No effort is made to reconcile the esimate from KF2 with the estimate from KFg, since this would typically cause a large jump in the combined state estimate at time t_h (on the order of the tracker bias) which could kick the image spot out of the slit. In fact, this is the reason that standard 3-axis control fails, and is avoided in the reconfigurable control concept.

3.3 Covariance Analysis

A single axis covariance analysis is given below, to characterize behavior along the slit dispersion direction.

Given desired time constant τ, the Kalman filter gains k_1 and k_2 (associated with a two-state observer of single-axis position and gyro hiss) are detuned as follows. Let $\omega_{kj} \triangleq (\frac{q_2}{r})^{\frac{1}{4}}$.

- If $\frac{1}{\tau} \leq \omega_{kf}$ then use complex roots: set $k_1 = \frac{2}{\tau}$ and $k_2 = \omega_{kf}^2$.
- If $\frac{1}{\tau} > \omega_{kf}$ then use repeated real roots: set $k_1 = \frac{2}{\tau}$ and $k_2 = \frac{1}{\tau^2}$.

The steady-state covariances associated with the detuned Kalman filters can be calculated as,

$$p_{11} = \frac{r(k_2^2 + k_1^2 k_2) + q_1 k_2 + q_2}{2k_1 k_2}, \quad p_{12} = \frac{rk_2^2 + q_2}{2k_2} \tag{9}$$

$$p_{12} = \frac{r(k_2^3 + k_1^2 k_2^2) + q_1 k_2^2 + q_2(k_2 + k_1^2)}{2k_1 k_2} \tag{lo}$$

$$P \triangleq E[ee^T] = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \tag{11}$$

where e = $[66, \delta b]^T$, $\delta\theta = 8 - \hat{\theta}$ is the error in the angular position estimate, $\delta b = b - \hat{b}$ is the error in the gyro bias estimate, and,

q_1 Gyro Angle Random Walk Covariance (rad^2/sec)
q_2 - Gyro Bias Instability Covariance (rad^2/sec^3)
r - Equivalent CT Tracker noise covariance (rad^2)
$r = \Delta\sigma_{nea}^2/N$
Δ - Tracker Sampling Period (see)
σ_{nea}- 'hacker NEA (per star, 1-sigma)
N - Number of stars on Tracker FOV

The pointing jitter after handoff can be calculated as,

$$\sigma_a^2 = \beta^2 \cdot p_{11}(\tau_2) \tag{12}$$

where $p_{11}(\tau_2)$ *is the* position estimation error covariance of KF2, and β = 206265 is a scale factor to convert radians to arcseconds. The quantity σ_a^2 will generally increase as KF2 is detuned further (i.e., as τ_2 is decreased).

The total pointing bias after handoff can be calculated as,

$$w_b^2 = \sigma_g^2(t) + \sigma_p^2 + \sigma_s^2 + \sigma_m^2 + \sigma_c^2 + \sigma_a^2 + \sigma_{KF2}^2(t) \tag{13}$$

where,

σ_g - Gyro Drift
σ_p - Body-to-Telescope Frame Misalignment error
σ_s - Gyro Scale Factor Error
σ_m - Gyro Frame Misalignment Error
σ_c - Steady-State Control Bias Error (after handoff)
σ_{KF2} - Bias from Kalman Filter KF2 settling
w_{sru} - Tracker bias change (over maneuver)

The jitter term σ_a^2 reappears in the bias expression (13) because at time $t = t_h$ one is clamping onto the random (rather than deterministic) process associated with KF2. The time-varying terms $\sigma_g^2(t)$ and $\sigma_{KF2}^2(t)$ dominate the expression for the bias (13) and deserve closer attention. The gyro drift is given by,

$$\sigma_g^2(t) = \beta^2 \cdot \left[\frac{q_2}{3}t^3 + p_{22}(\tau_1)t^2 + (2p_{12}(\tau_1) + q_1)t + p_{11}(\tau_1)\right] \tag{14}$$

where $p_{ij}(\tau_1)$ are steady-state covariances from the detuned filter KF1. As shown in Figure 2 the gyro drift increases monotonically with time after $t = O$. The settling bias of KF2 is given by,

$$\sigma_{KF2}(t) = w_{sru}e^{-(t-t_a)/\tau_2} \tag{15}$$

where τ_2 is the time constant (by design) associated with KF2. This is the error associated with clamping onto the filter KF2 before it has completely settled, As shown in Figure 2 the settling bias decreases monotonically with time after $t = t_a$.

3.4 Application to SIRTF IRS

The reconfigurable control concept is applied to the SIRTF telescope in support of the IRS payload. Parameters associated with a candidate SIRTF pointing control design are given as q_1 = 3 .3846*e* - 15 rad^2/sec, q_2 = 7.3451e -21 rad^2/sec^3, r = 1.0581*e* -12 rad^2, $\sigma_p = \sigma_s = \sigma_m = \sigma_c$ == .1 *arcsec*, w_{sru} = *.4 arcsec*, where a 30 arcmin maneuver has been assumed, Parameters relevant to the IRS payload are given as [4] A_2 = .13 $(arcsec)^{-2}$, d = .07, α = .05.

Equation (12) for σ_a^2 and (13) for w_b^2 are substituted into (6) to give the quantity $\tilde{x}_{.95}$ for $t \geq t_a$, which is plotted in Figure 3 for different values of τ_1 = 10,20,30 and τ_2 = 10,20,30. If the handoff is timed to catch the minimum of each curve, it is seen that the desired value of d = .07 can be satisfied with any one of several possible designs. For example one reasonable design would be τ_1 = 20,72 = 20 which requires optimal handoff at $t = t_a + 30$ seconds, and achieves a performance better than d = .06. Without reconfiguration, a 3-axis controller for the same example would perform no better than *d = .12,* and would have additions drift terms which have not been analyzed here.

4 Conclusions

A pointing control performance criteria has been established in support of high resolution space spectroscopy. The requirement, given by (7), simultaneously constrains both the pointing bias and jitter to ensure that the flux offset is small in the sense that it is less than a specified fractional error d with at least $(1 - \alpha)$ x 100% percent confidence. Calculations indicated that these pointing requirements would be difficult to meet using standard 3-axis spacecraft control primarily due to a tight pointing hiss requirement.

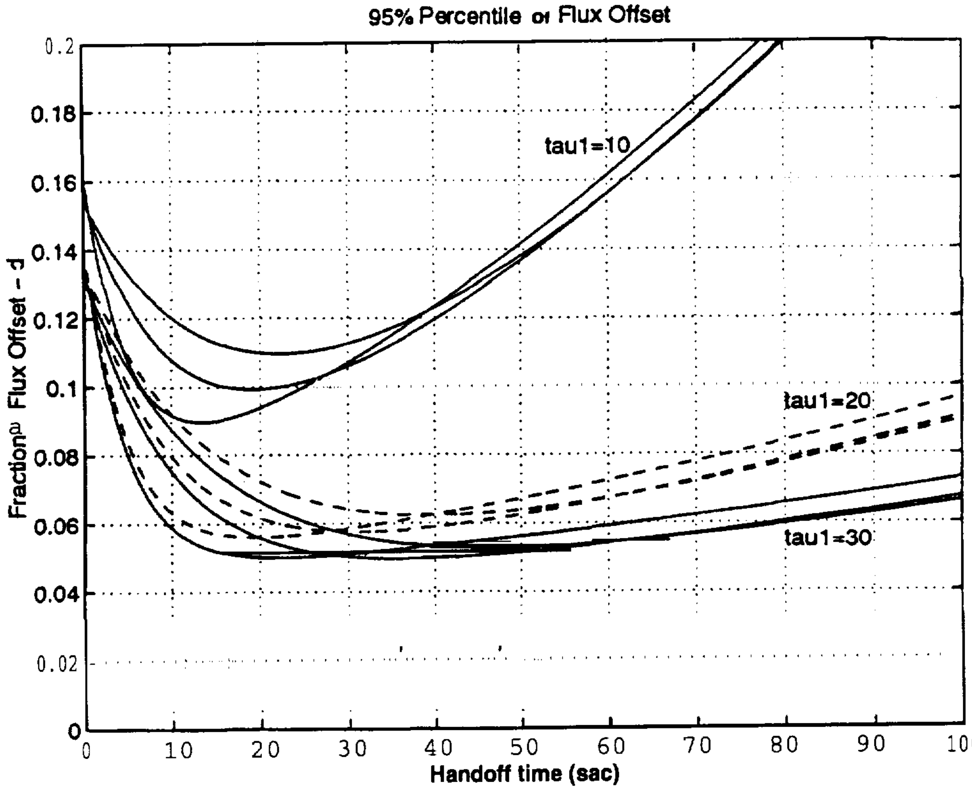


Figure 3: Optimal Handoff Timing and Performance

In order to satisfy the performance requirement, a reconfigurable control concept was proposed which avoids to a large extent the contribution of the bias error from the star tracker. The effectiveness of the control reconfiguration approach was demonstrated on the SIRTF pointing system in support of the IRS payload. Results indicate that by proper choice of falter detuning and optimized handoff timing, the flux offset can be held (with ,95 probability) to within d = .06 of the ideal flux. This contrasts with d = .12 for the 3-axis control design, and results in significantly improved high-resolution science capability.

Acknowledgements

The authors would like to thank Fernando Tolivar of JPL for several technical discussions. This research was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration,

References

[1] D.S. Bayard and T, Kia, "IRS Pointing Requirements for SIRTF Under 3-Axis Control," JPL Internal Document, Engineering Memorandum EM-3454-97-008, November 13, 1996.

[2] A. Papoulis, *Probability, Random Variables and Stochastic Processes.*, Second Edition, McGraw Hill, New York, 1984.

[3] S.W. Sirlin, A.M. San Martin, " A New Definition of Pointing Stability," JPL Internal Document, Engineering Memorandum EM 343-1189, March 6, 1990.

[4] J. Van Cleve, "SIRTF Pointing Requirements Derived from Slit Transmission in the Diffraction Limit," Report, Dept. of Astronomy, Cornell University, Ithaca NY, May 14, 1991.

[5] J. Van Cleve, "Jitter-Drift Pointing Requirements for SIRTF-IRS," Report, Dept. of Astronomy, Cornell University, Ithaca NY, July 10, 1995.

[6] J. Van Cleve, "Accuracy and Reconstruction Models," Report, Dept. of Astronomy, Cornell University, Ithaca NY, September 10, 1996.

URC97166

/// //
5/65/52

Dexterity-enhanced Telerobotic Microsurgery

Steve Charles

MicroDexterity Systems, Inc. and Charles Retina Institute, Memphis, TN

Hari Das, Timothy ohm, Curtis Boswell, Guillermo Rodriguez, Robert Steele

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA.

Dan Istrate

California Institute of Technology, Pasadena, CA.

1 Introduction

The work reported in this paper is the result of a collaboration between researchers at the Jet Propulsion Laboratory and Steve Charles, MD, a vitreo-retinal surgeon. The Robot Assisted MicroSurgery (RAMS) telerobotic workstation developed at JPL [9] is a prototype of a system that will be completely under the manual control of a surgeon. The system, shown on Figure 1, has a slave robot that will hold surgical instruments. The slave robot motions replicate in six degrees of freedom those of the surgeon's hand measured

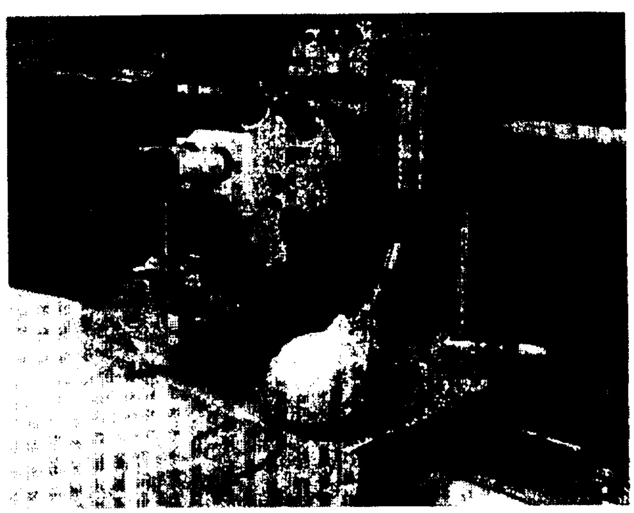


Figure 1: RAMS telerobot system.

using a master input device with a surgical instrument, shaped handle. The surgeon commands motions for the instrument by moving the handle in the desired trajectories. The trajectories are measured, filtered, and scaled down then used to drive the slave robot,.

We present the details of this telerobotic system by first giving an overview of the subsystems and their interactions in the following section then present det ails in subsequent subsections divided according to subsystem. This paper concludes wit h a description of a recent demonstration of a simulat ed microsurgery procedure performed at JPL.

2 System description

Figure 2 shows an overview of the hardware components of the RAMS telerobotic system.

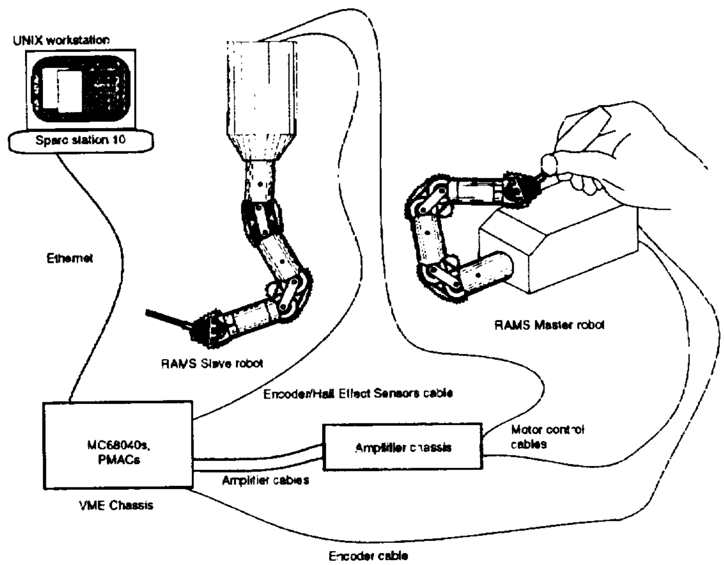


Figure 2: RAMS slave robot system.

Components of the RAMS system have been categorized into four subsystems. They arc the mechanical subs ystem, the electronics subsystem, the servo-control subsystem and the high-level software subs yst em. The mechanical subsystem consists of a master input device and a slave robot arm with associated motors, encoders, gears, cables, pulleys and linkages that cause the tip of t he robot to move under computer control and to measure the surgeon's hand motions precisely. The electronics subsystem consists of the motor amplifiers, a safety electronics circuit and relays within the amplifier box shown on Figure 2. These elements of the subsystcm ensure that a number of error conditions are handled gracefully. The servo-control subsystem is implement ed in hardware and software. The relevant hardware parts of the subsystem are the servo-control boards and the computational processor boards. Servo-control software functions include setting-up the control parameters and running the servo-loop on the servo-cent rol board to control the six motors, implementing the communication between the computation and servo-control boards, initializing the servo-control system and communicating with the electronics subsystem and communicant ing with the high-level software subsystem. The high-level software subsystem interfaces with a user, controls initialization of the system software and hardware, implements a number of demonstration modes of robot cent rol and computes both the forward and inverse kinematics. A drawing of the interaction between the subs yst ems oft he RAMS slave robot, is shown on Figure 3.

2.1 Mechanical subsystem

The RAMS slave manipulator is a six degrees-of-freedom tendon-driven robotic arm designed to be compact yet exhibit very precise 10 micron relative positioning capability as well as maintain a very high work volume, Physically, the arm measures 2.5 cm. in diameter and is 25.0 cm. long from its base to tip. It is mounted to a cylindrical base housing which measures 12 cm. in diameter by 18 cm long that contains all of the drives that actuate the arm. A photograph of the arm appears on Figure 4. The joints of the arm are a torso joint rotating about an axis aligned with the base axis and positioned at the point the arm emerges from its base, a shoulder joint rotating about two axes that are in the same plane and perpendicular to the preceding links, an elbow joint that also rotates about two axes that are in the same plane and perpendicular to the preceding links, and a wrist with pitch, yaw and roll joints.

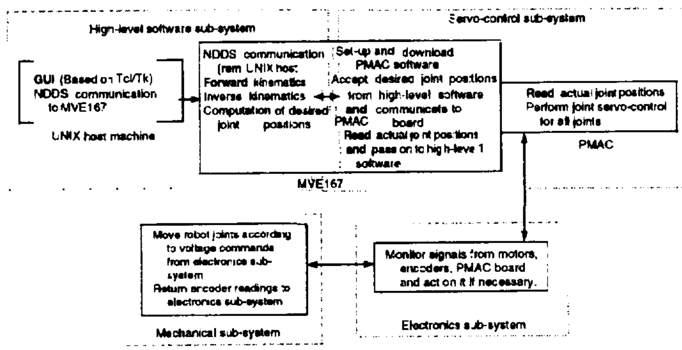


Figure 3: Sub-systems of the RAMS telerobot system.

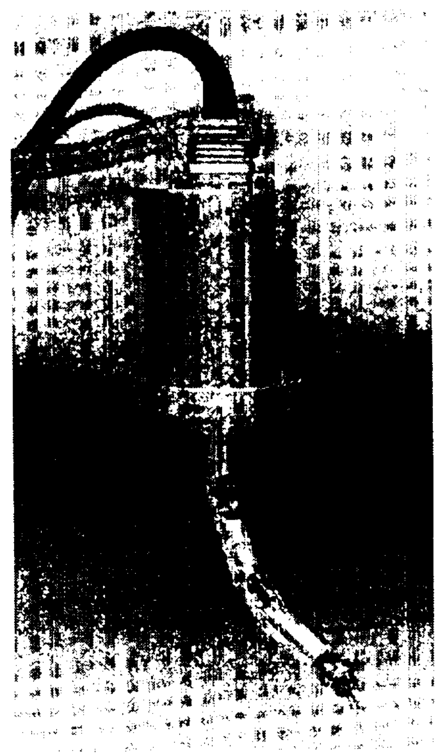


Figure 4: RAMS slave robot.

The master device, cinematically similar to the slave robot, also has six tendon driven joints. It is 2.5 cm. in diameter and 25 cm. long. Its base houses high-resolution optical encoders requiring a larger volume - a box of size 10.8 cm by 18.4 cm by 23.5 cm. Gear transformation ratios in the master arm are reduced to allow backdrivability. A photograph of the master input device *is* shown on Figure 5.

The slave wrist design (based on the kinematics of the Rosheim OMNI-WRIST [8]) utilizes a dual universal joint to give a three degrees-of-freedom, singularity free, mechanically decoupled joint that operates in a full hemisphere of motion (up to 90 degrees in any direction). The master wrist design uses a universal joint to transmit rotation motion through the joint while allowing pitch and yaw motions about the joint resulting in singularity free motion over a smaller range of motion in three degrees-of-freedom. The fourth and fift h axes of the master and slave robots arc unique joints that rotate about 2 axes and allow passage of cables to pass through the joint for actuating the succeeding joints without affecting their cable lengths. The sixth axes are torso joints which simply rot ates the manipulators relative to their base housing, For t hc slave robot

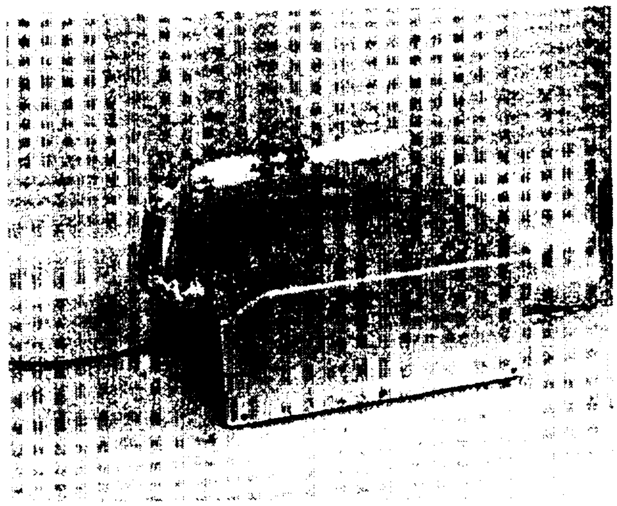


Figure 5: RAMS Master input device.

the torso range of motion is 330 degrees while on the master it is 30 degrees.

2.2 Electronics subsystem

The RAMS electronics subsystem design includes off the shelf and custom designed electronics. Figure 6 shows a layout of its general components.

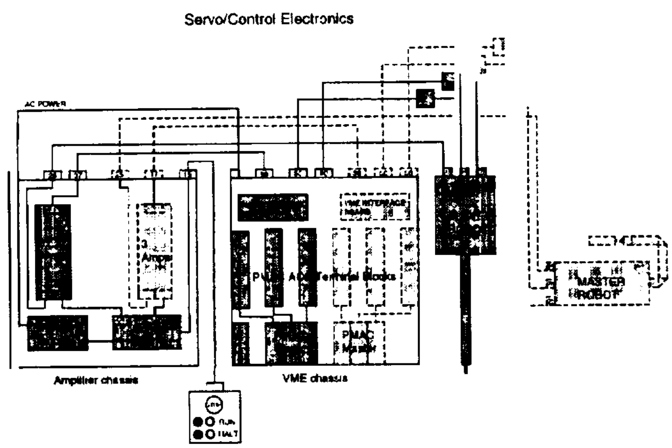


Figure 6: Electronics components and cabling.

Components of the electronics subsystem are a VME chassis, an amplifier chassis and safety electronics. The VME chassis houses the VME backplane and two Motorola MVME-167 computer boards used for high level system control. The VME chassis also contains the PMAC servo control cards and six supporting interface modules, power supplies (+/-15v) and a cable interface board. The VME chassis front panel contains main power control (AC) for the system. The rear panel provides access to the cent rol computers serial communications port (RS-232). All components above are off the shelf items except the cable interface board.

2.3 Servo-control subsystem

The RAMS servo-control system is implemented on processor boards and servo-control boards installed in a VME chassis. Two Motorola MVME-167 boards, named *Proc0* and *Proc1*, are installed on the VME chassis and run under the VxWorks operating system, *Proc0* performs kinematic, communication and high-level control functions. These functions are described in the High Level Software Architecture Section. Calls to subroutines that read and set joint angle positions of the robot arc made from the high-level real-time software on *Proc0*. These routines, through shared memory implemented between *Proc0* and *Proc1*, provide setpoints and read current joint angles of the robot. *Proc1*, in turn, passes the setpoints for controlling the robot to the servo control board and retrieves the joint angles measured by the servo-control board. The servo level control system uses the PMAC-VME board by Delta Tau.

Communication between *Proc0, Proc1* and the PMAC-VME boards is through shared memory. The PMAC board has a large variety of features for motor control, with a customer base largely from industrial installations. The key features used for control of the RAMS robot include:

2.4 High-level software subsystem

There arc a number of components to the high-level software for the RAMS slave robot. A drawing of the parts of the software in shown on Figure 7. Embedded in the computational blocks of the real-time

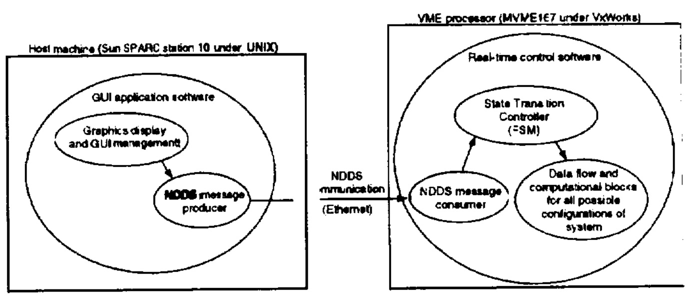


Figure 7: Parts of the high-level software.

control software are the kinematic control algorithms. They are based on algorithms developed at JPL [6],[7] for the unique geometry of the robot. The demonstration of different control modes of the robot was implemented using a software development tool for real-time systems called Control Shell [3], [4]. Handling of operator commands in the real-time software, transitions between states of control, changes in data flow due to transitions of states in the software and the algorithms executed within computation blocks. The user specifies the control modes of the system through a graphic user interface (GUI) implemented with Tcl/Tk [2]. Commands entered into the GUI are transmitted over an Ethernet connection and are received on the real-time software side of the system. The message passing between the 2 parts of the software system uses NDDS [5]. A *producer* part creates the messages and broadcasts them from the GUI part of the system and a consumer part receives the messages and processes them.

3 Simulated Surgery

In September of 1996, a demonstration of a simulated eye microsurgery procedure was successfully conducted using the RAMS telerobotic system. The procedure demonstrated was the removal of a microscopic 0.015 inch diameter particle from a simulated eyeball.

Features added to the RAMS system to enable successful performance of the eye surgery demonstration were foot switch operated indexed motion, a surgical instrument mounted on the slave robot tip and a pivoting shared control algorithm to automatically compensate for pitch and yaw orientation of the surgical

instrument while the operator cent rols the x-, y-, z- and roll motions of t hc instrument. Figure 8 shows the RAMS system as seen performing the simulated eye microsurgery procedure.

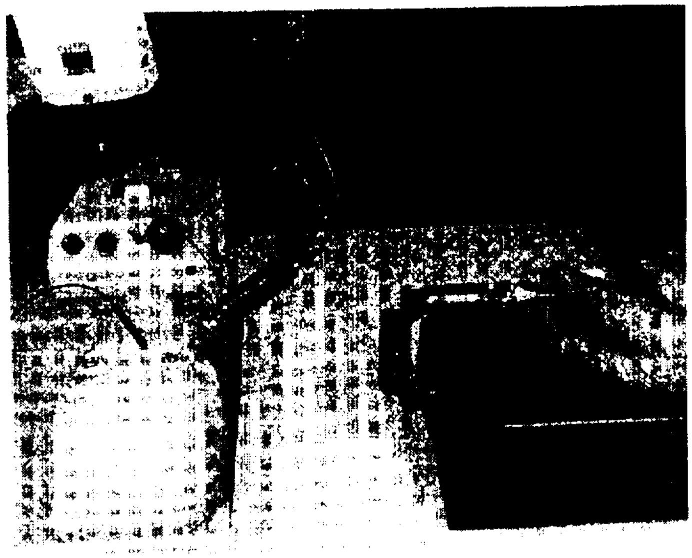


Figure 8: Performing the eye surgery demonstration.

In the next year, the RAMS system will be upgraded to implement force feedback to the master arm from force sensors mounted on the slave robot. In addition, experiments will be conducted to characterize the performance of the system as compared to direct manual manipulation in simulations of microsurgical tasks.

References

[1] Charles, S. "Dexterity Enhancement for Surgery", in Computer Integrated Surgery: Technology and Clinical Applications, ed. R. H. Taylor, S. Lavalle, G. Burdea, R. Mosges, MIT Press, Cambridge, MA 1996.

[2] Ousterhout, J. K. "Tcl and the Tk Toolkit", Addison Wesley, Reading Mass. 1994.

[3] Real-time Innovations, Inc., 'Control Shell Programmer's Reference Manual Vol. 1", Sunnyvale, CA, 1995.

[4] Real-time Innovations, Inc., 'Control Shell Programmer's Reference Manual Vol. 2", Sunnyvale, CA, 1995.

[5] Real-time Innovations, Inc., "NDDS Programmer's Reference Manual", Sunnyvale, CA, 1995,

[6] Rodriguez, G., K. Kreutz, and A. Jain, "A Spatial Operator Algebra for Multibody System Dynamics," The Journal of the Astronautical Sciences, Vol. 40, No. 1, pp. 27-50, January-March 1992.

[7] Rodriguez, G., "Kalman Filtering, Smoothing, and Recursive Robot Arm Forward and Inverse Dynamics," IEEE Transactions on Robotics and Automation, Vol. 3, pp. 624-639, Dec. 1987.

[8] Rosheim, M. E., "Robot Wrist Actuators", John Wiley & Sons., New York, 1989.

[9] Schenker, P., Das, H., and Ohm, T. ".4 new robot for high dexterity microsurgery" Proceedings of the First International Conference, CVRMed '95, Nice, France April, 1995. also in Computer Vision, Virtual Reality and Robotics in Medicine, Lecture Notes in Computer Science, Ed. Nicholas Ayache, Springer-Verlag, Berlin 1995.

[10] Williams III, R. L., "Forward and Inverse Kinematics of Double Universal Joint Robot Wrists," Space operations Applications and Research (SOAR) Symposium, Albuquerque, NM, June 26-28, 1990,

Acknowledgment

This work was carried out at the Jet Propulsion Laboratory under contract with the National Aeronautics and Space Administration. The authors affiliated with JPL are in the Automation and Control Section, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109. Steve Charles, MD is the CEO of MicroDexterity Systems, Inc.

~~16~~ 6/62

URC97167

Mars Surveyor '98 Lander MVACS Robotic Arm Control System Design Concepts

Robert G. Bonitz

Telerobotics Research and Applications Group

Jet Propulsion Laboratory, California Institute of Technology

4800 Oak Grove Drive, Pasadena, CA 91109-8099

Abstract

This paper describes the control system design concepts for the Mars Volatiles and Climate Surveyor (MVACS) Robotic Arm which supports the scientific investigations to be conducted as part of the Mars Surveyor '98 Lander project. Solutions are presented to some of the problems encountered in this demanding space application with its tight constraints on mass, power, volume, and computing resources. Problems addressed include 4-DOF forward and inverse kinematics, trajectory planning to minimize potential impact damage, joint drive train protection, Lander tilt prevention, hardware fault monitoring, and collision avoidance.

1 Introduction

In January of 1999 NASA will launch the Mars Surveyor '98 Lander spacecraft with the Mars Volatiles and Climate Surveyor (MVACS) integrated science payload as part of the Mars Exploration Program [1]. The target landing site is at 71° S latitude late during the Martian southern spring season. The goal of the MVACS mission is to conduct scientific investigations to characterize the surface environment, search for near-surface ground ice, determine the quantity of CO_2 and H_2O in the soil, determine the abundance of volatile-bearing minerals in the soil, and search for climate records in the form of fine-scale layering in near-surface materials.

The major elements of the MVACS† science payload consist of (see Figure 1):

1. A Surface Stereo Imager (SSI) supplied by the University of Arizona (UA) to characterize the general environment;
2. A Robotic Arm Camera (RAC) supplied by (UA) and the Max-Planck Institute for Aeronomy to provide close-range surface and subsurface images;
3. A Thermal and Evolved Gas Analyzer (TEGA) supplied by UA to determine concentrations of ices, adsorbed volatiles, and volatile-bearing minerals in soil samples;
4. A Meteorology Package (MET) and Soil Temperature Probe (STP) supplied by the Jet Propulsion Laboratory (JPL) to measure pressure, wind speed, and temperature;
5. A two-meter 4-DOF Robotic Arm (RA) supplied by JPL to dig trenches, acquire soil samples, position the RAC, position the Soil Temperature Probe, and deposit soil samples in the TEGA.

The Lander spacecraft is being built by Lockheed Martin Astronautics. The Principal Investigator is Professor David Paige of UCLA with project management provided by JPL.

Since this is a space application, there are tight constraints on mass, power, volume, and computing resources. The Robotic Arm is allocated a mass of 3.5Kg and a 10 watt average (20W peak) power budget.

†http://mvacs.ess.ucla.edu/

Table 1: MVACS Robotic Arm Task Commands

Name	Description
ra_acquire_sample(x, y, z,type)	Acquire a sample at location x, y, z.
ra_dig_trench(x, y, z ,depth ,length, width, trench-angle, dig_angle)	Dig a trench starting at x, y, z with the specified parameters.
ra_tega_dump()	Dump the sample into the current TEGA port.

The Robotic Arm control software will run on a RAD6000 running at 5MHz in a multitasking environment under the VxWorks[‡] operating system. Additionally, the integrated instrument payload makes maximum use of subsystem designs that have been previously qualified on the Mars Pathfinder program to reduce cost and mission risk. The purpose of this paper is to present the design concepts of the Robotic Arm Control System delineating some of the problems encountered in this application.

2 Robotic Arm System Design

The Robotic Arm system consists of a lightweight 4-DOF manipulator, Arm Control Electronics (ACE) to drive the motors and read in sensor data, and the control software to provide command expansion, trajectory generation, control compensation, and to monitor the sensor data. Salient features of the manipulator are:

1. 4-DOF back-hoe design sufficient to achieve digging, sample acquisition, and RAC and STP positioning;
2. Low-mass graphite-epoxy tubular links;
3. Joint actuators consisting of low-power motors with two-stage speed reduction (planetary gear plus harmonic drive) to achieve low-mass with high-torque output;
4. Capability of exerting up to 170N force at the end-effecter for digging with up to 800N for ripping with the tines in selected configurations;
5. Non-backdriveable joints (via motor detents) which can be selectively shut off to conserve energy during select ed manipulator motions;
6. Actuator heaters to maintain adequate temperature in the cold Martian environment;
7. Joint encoders and potentiometers to achieve 1cm absolute (0.5cm relative) positioning accuracy.
8. End-effecter tools including the RAC, STP, scoop for digging, and tines for ripping;
9. A payload capacity of 5Kg at full-arm extension.

The Arm Control Electronics consists of circuitry to read in data from the encoders and pots, provide current drive to the motors, and set motor voltage levels, and an 8032 microprocessor to execute low-level joint set-point commands received from, and to pass sensor readings to, the control software.

The control software receives high-level task commands from the ground via the Lander and expands them into Cartesian and joint-motion commands. Arm joint trajectories are computed for each motion command and each via point is sent to the ACE for execution. The control software also monitors the sensor data for collision avoidance and hardware fault detect ion. An example of a high-level task command is to dig a trench with a specified depth, length and starting location. The control software expands the dig trench command received from the ground into a sequence of Cartesian motion commands which are executed in turn. This provides ease of use of the Robotic Arm by the ground operator and conserves uplink bandwidth. To provide maximum flexibility, however, the ground operator may also send both Cartesian and joint-motion commands. A sample of some of the available task-level and low-level commands are given in Tables 1 and 2, respectively. The MVACS Robotic Arm control system is depicted in Figure 2.

[‡]Wind River Systems, Inc., 1010 Atlantic Ave., Alameda, CA

Table 2: MVACS Robotic Arm Move Commands

Name	Description
ra_move_cartesian (z, y, z, θ | d,ref)	Move the tool to the designated position. ref=absolute | relative | tool
ra_move_joint(jt1,jt2,jt3,jt4,ref)	Move the joints the designated number of radians or seconds. ref = absolute | relative | time

3 Kinematics

Since the MVACS Robotic Arm has only four degrees of freedom, complete specification of the Cartesian position and orientation of the cur-rent tool frame is not possible. One must choose which four degrees of freedom out of the six degrees that completely specify the position and orientation of the tool frame. For ease of operation by the user, the 4-DOF representation was chosen as X = $[x\ y\ z\ 0]^T$ where $[x\ y\ z]^T$ is the Cartesian position of the origin of the tool frame in the world-coordinate system and θ is the angle the tool approach vector makes with the x-y plane (parallel to the Lander deck). Frame assignments for the manipulator are given in Figure 3. Each of the tool frame approach (z) vectors is orthogonal to the rotation axis of the wrist joint so that the desired orientation is achievable. One exception is the RAC which is attached to the forearm and, thus, there are only three degrees of freedom. For the RAC, the user only specifies the desired position of the tool frame and not orientation.

The MVACS 4-DOF manipulator presents a unusual inverse kinematics problem. A typical method for solving the inverse kinematics for a 6-D OF manipulator is to compute the desired location of the wrist frame from the specified position and orientation of the tool frame and then use geometric methods to solve for the first three joints followed by the last, three joints [2]. For a 4-DOF manipulator with only one orientation angle specified, the above method fails since one does not have the full rot ation matrix of the tool frame and cannot compute the wrist frame location.

In the MVACS case, however, it is possible to solve directly for joint 1 from the specified tool-frame position using geometric techniques since the last three joint axes are parallel and orthogonal to the axis of joint 1. One can then solve for the wrist position from

$$^{w}p_4 = {}^{w}p_{tool} - {}^{w}R_4{}^{4}p_{tool} \tag{1}$$

where $^{w}p_4$ is the wrist location, $^{w}R_4$ is the rotation matrix of the wrist, $^{4}p_{tool}$ is the vector from the wrist-frame origin to the tool-frame origin expressed in the wrist frame, and $^{w}p_{tool}$ is the desired position of the tool in world-frame coordinates. From the forward kinematics, we get

$$^{w}R_4 == \begin{bmatrix} c_1 c_{234} & -s_1 & c_1 s_{234} \\ s_1 c_{234} & c_1 & s_1 s_{234} \\ -s_{234} & 0 & c_{234} \end{bmatrix} \tag{2}$$

where $s_i = \sin(\sum q_i), c_i = \cos(\sum q_i), q_i$ is the ith joint angle, and $q_2+q_3+q_4 = {}^{w}\theta_{tool} - {}^{4}\theta_{tool} + \pi/2$. $^{w}\theta_{tool}$ is the specified tool-frame orientation angle and $^{4}\theta_{tool}$ is the angle from the wrist-frame z axis to the tool-frame z axis about the wrist-frame y axis. Now that the wrist position is know, one can use standard geometric techniques to solve for remaining joint angles.

Since the RAC is affixed to the forearm and not the end effecter, it presents an unusual inverse kinematics problem as well since we cannot, specify both its location and orientation. Furthermore, it would difficult for the operator to specify the location of the RAC given that the location of the target to be imaged is known. To make it easy on the operator when the tool is set to the RAC, Cartesian motion commands are interpret ed as $[x\ y\ z]^T$ being the location of the target to be imaged and the fourth element, d, as the distance from the RAC to the target. To solve the inverse kinematics, the following method is used:

1. Redefine the forearm to be the line from the elbow joint to the target when the RAC approach vector points at the target and compute its length from d and the arm geometry;

2. Set ${}^{w}p_4$ to be the target location;

3. Solve for the three joint angles using standard geometric techniques;

4. Compute an adjustment to q_3 from d and the arm geometry necessary because of step 1.

4 Trajectory Planning

The Robotic Arm has the capability to execute both Cartesian and joint motion commands. Cartesian mot ion can be expressed as absolute or relative motion in the world frame or as tool-frame motion. Furthermore, straight-line or joint-interpolated motion can be specified. Joint-space motion commands can be given as absolute joint angles, as relative joint angles with respect to the current joint positions, or as timed mot ion. In order to prevent the arm from moving too far and potentially damaging Lander hardware in the event of loss of communication between the ACE and the control software, a sequence of via points is computed for any motion command and only after the arm has reached the current via point is the next via point sent to the ACE. Thus, trajectories are generated as a sequence of points in space only and not in time. The one exception is the timed joint-motion commands in which the operator can command each joint to move for a specified amount of time in which case the generated trajectory is a sequence of via points in time only and not in space.

The manipulator joints have high gear ratios and joint velocities are slow (on the order of tenths of a degree per second). During spatial arm motion at least one joint is moving at its maximum velocity. The remaining joint velocities are appropriately scaled to achieve coordinated motion. For timed joint motions, the operator can set the motor voltages as desired which will determine the joint velocities.

For Cartesian via sequence generation, the bounded-deviation joint-path method in [3] is used. This algorithm is basically a recursive bisector method for finding the via points. The straight-line Cartesian position at the midpoint is compared to the Cartesian position calculated at the midpoint oft he corresponding joint space. If the error exceeds a specified amount (set to smaller than the required accuracy), the path is divided into two segments and the algorithm executed for each segment. The process cent inues until the error limit is satisfied.

Since we require via sequencing for both spatial and timed joint-motion commands as well, a similar approach is taken to generate the via sequence. The algorithm is as follows:

1. Store the current position as the first via point (for timed moves the starting via is set to zero);

2. Compute the joint-motion midpoint;

3. If the amount of the joint motion exceeds the specified limit, execute steps 2-4 for the two segments (start, point, midpoint) and (midpoint, end point);

4. Else store the end point as the next via point.

The limit and the amount of joint motion is specified as a single metric and is represented by $||q||$ where q is the joint vector in radians or seconds.

5 Control System

In addition to task expansion and trajectory generation, the Robotic Arm control software also monitors progress of the arm towards its commanded via point as well as joint torque and Lander tilt-moment limits. Compensation is proportional control with sufficient gain such that at least one motor voltage is saturated to achieve maximum velocity until the arm approaches the set point. The remaining motor volt ages are scaled accordingly.

Due to the high gear ratios, it is possible for the motors to exert enough torque to exceed the specified torque limits of the harmonic drives. During free-space arm motion, the load on the joints will not exceed the torque limits. However, during digging it is possible to exceed the limits if the motion of the end effecter is impeded by rock, ice, or hard soil. Furthermore! under the preceding condition in certain arm configurations, it is also possible for the arm to exert enough force and torque on the Lander to tilt it; viz., high enough to

lift one or more footpads off the ground. It is critical that, joint torques be monitored to assure that damage or tilting not occur.

The MVACS Robotic Arm is not equipped with a force-torque sensor, but the motor currents are sensed so it is possible to estimate joint torques whenever the motors are on using the motor torque constant which is nearly linear. Once the full set of joint torques is computed, the reaction force and torque at the Lander deck can be computed and used along with the Lander geometry to compute the tilt moments about the lines between the footpads. The joint torques can be compared to the harmonic drive torque limits and the tilt moments can be compared to the tilt-moment limits to assure that damage or tilting does not occur. If either of the limits are being approached, the planned trajectory is modified by computing a delta to the trajectory which reduces dig depth by using a PI filter in the force feedback path [4]. If the limits are exceeded, the arm is stopped and the situation assessed by imaging the scene with the SSI and evaluating the downlinked images along with the engineering data.

6 Fault Monitoring and Collision Avoidance

It is critical that the Robotic Arm operate safely during the execution of its assigned tasks so as not to damage itself or any of the Lander hardware. Each time through the control loop, sensor data is analyzed and an assessment made as to whether any hardware failures have occurred. Available sensor data include joint posit ions from both encoders pot volt ages, motor currents, joint temperatures, power supply status, and A/D reference voltages. Pot ential hardware faults include failures of t he sensors, motors, power supply, or voltage reference. The position of the joints as determined from both the encoders and pot voltages is also assessed and if the difference exceeds a specified limit, the arm is recalibrated. The calibration process consists of driving the arm against the joint limits and reloading the encoder counters. During normal operation the encoders are used as the primary joint-position sensor. In the event of an encoder failure, the arm can be commanded to use the pot voltage as the primary position sensor with some degradation of positioning accuracy.

To prevent collisions of the arm with the Lander hardware, all command sequences will be verified on the ground by simulation to assure t hat kinematic stay out, zones are not violated. The Telegrip§ robotic system simulation tool is used to simulate the arm motion. In addition, on board collision detection algorithms are executed both during path planning and while the arm is in motion. The MVACS Robotic Arm uses the obstacle detection scheme developed for the NASA Ranger Telerobotic Flight Experiment [5] adapted for the MVACS environment. This scheme is model based and uses simple object models and distance computation algorithms making it suitable for the MVACS real-time application. In the event of detection of an imminent collision, the arm is stopped and the situation assessed by imaging the scene with the SSI and evaluating the downlinked images.

7 Conclusion

Design concepts of the Mars Surveyor '98 Lander MVACS Robotic Arm Control System were presented describing some of the unique problems encountered in this demanding space application with its tight constraints on mass, power, volume, and computing resources. Solutions to the inverse kinematics problem, trajectory generation, damage and tilt prevention, hardware fault monitoring, and collision avoidance were described.

References

[1] D.A. Paige, et al, "Mars Volatilcs and Climate Surveyor (MVACS), Integrated Payload Proposal for the Mars Surveyor Program '98 Lander," 1995. NASA AO No. 95-0 SS-3.

[2] K. Fu, R. Gonzales, and C. Lee, *Robotics:* Control, Sensing, Vision, *and Intelligence.* CAD/CAM, Robotics, and Computer Vision, McGraw-Hill, 1987.

[3] R. Taylor, "Planning and execution of straight line manipulator trajectories," *IBM Journal of Research and Development,* vol. 23, no. 4, pp. 424–436, 1979.

§Deneb Robotics, 3285 Lapeer Road West, Auburn Hills, MI

[4] R. Bonitz and T. Hsia, "Robust internal force-tracking impedance control for coordinated multi-arm manipulation - theory and experiments ," in *6th International* Symposium *on Robotics and Manufacturing, 2nd World Automation Congress,* (Montpelier, France), May 1996.

[5] B. Bon and H. Seraji, "Obstacle detection for the Ranger Telerobotic Flight Experiment ," in *6th International* Symposium *on Robotics and* Manufacturing, *2nd World Automation Congress,* (Montpellier, France), May 1996.

Figure 1: Mars '98 Lander with MVACS

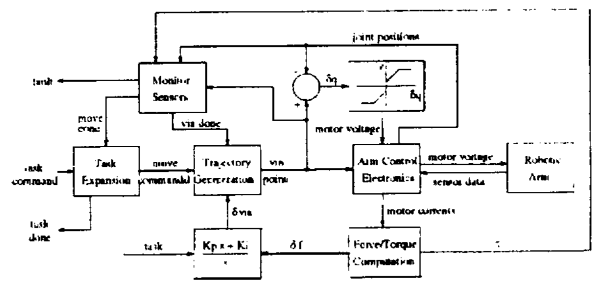


Figure 2: MVACS Robotic Arm Controller

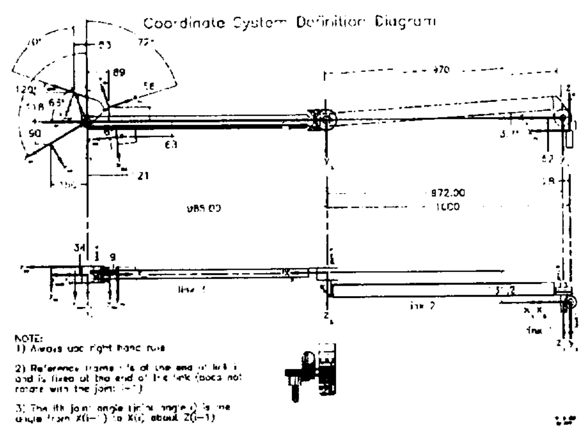


Figure 3: Manipulator Frame Assignments

URC97168

5/67/57
NRS 1147

The Lagoon Nebula M8

ANTHONY WILLIAMS

South Carolina State University

ABSTRACT

The Lagoon Nebula (M8) is a well known H II region in the constellation of Sagittarius. It is one of the most studied objects in the Interstellar Medium (ISM) and has been examined at all wavelengths from the radio to the gamma ray region. This study will concentrate on the optical region using CCD images taken through filters centered at select wavelengths. These wavelengths are important to understanding the physics of the nebula and include emission lines of the following ions: 0^{++} at $\lambda\lambda$4363,5007, N^{+} at $\lambda\lambda$5755,6584, S^{+} at $\lambda\lambda$6717,6731 and H^{+} at $\lambda\lambda$4861,6563. The first two sets of lines are used to determine the electron temperature (T_e) of the nebula, the next pair is used to calculate the electron density (N_e) and the final pair is used to determine the amount of light which is scattered due to dust in the nebula.

The CCD images used in this study were obtained at San Pedro Matir Observatory in Baja, Mexico, using the 2.12 meter telescope. The images were reduced using the Image Reduction Analysis Facility (IRAF). Corrections to the raw data included bias subtraction, dark subtraction, cosmic ray removal and corrections for flat field differences. Additional steps were necessary to produce the final images, which are T_e and N_e maps of the nebula. These steps will be discussed. The maps show variations in temperature and density on a pixel-by-pixel level.

The 0^{++} images clearly show high ionization regions in the nebula near the ionizing stars, as expected, Some of the most dramatic variations can be seen in the sulfur images. Clumps, ridges and arcs of high density material are clear and well defined. These maps are used to better understand the physical conditions in M8 and the chemical abundances in the nebula.

URC97169

Temperature and Density Variations in Galactic Nebulae

DONALD K. WALTER

South Carolina State University

ABSTRACT

The results from spectrophotometric studies of several galactic emission nebulae are discussed. Details of the spatial variation of the electron temperature (T_e) and density (N_e) are given for the Orion Nebula (M42), with a less in-depth discussion of the Lagoon Nebula (M8) and the Ring Nebula (M57), both of which are covered elsewhere in these proceedings by Anthony Williams and Latongia Groce respectively.

The data set consists of ground based longslit spectra and CCD imagery taken through narrow-band interference filters centered on important diagnostic lines of 0^{++}, N^{+} and S^{+} and H^{+}. These seeing-limited (< 2") images show significant spat ial variations in T_e and N_e on scales as small as 0.005 parsec. Global trends and local variations of these diagnostics will be examined.

For the Orion Nebula it has been known for some time that N_e determined from the S^{+} double at $\lambda\lambda$6717,31 decreases with distance from the stellar ionizing source, θ^1 C Ori. Our data show a similar trend for the Cl^{-+} ion using the doublet at $\lambda\lambda$5518,5538. These results are consistent with models of H II regions in the champagne phase of evolution.

The existence of temperature gradients in the Orion Nebula has been debated for over twenty years. We have used one, self-consistent set of data to convincingly show that these gradients do exist and are different for each ion. The 0^{++} and S^{+} data can be fit with a gradients which increase with distance from θ^1 C Ori, but do show local deviations from the fits. The value for T_e of the N^{+} ion shows a sharp drop in temperature out to a distance of 100 arcseconds from θ^1 C Ori and then a very gradual decrease to a low of 8500 K at a distance of 300 arcseconds. These results must be taken into account by future models of H II regions.

Since nebular abundances are normally determined from forbidden-line ratios, and these lines are very sensitive to temperature ($\propto \exp[-1/kT]$) a difference of only 500-1000 K can result in abundance calculations which differ by a factor of two or more. For nebular abundance calculations, the importance of using local, on-the-spot values for T_e and N_e rather than global values is demonstrated.

URC97170

5169/53

Can the Ionospheric Disturbances of Gamma-Ray Bursts be Detected Using the VLF Method?

J. Brodney Fitzgerald, William Wright, Philip George, Hezzick Doyle,
Charles H. McGruder, III, Johnny Jennings

Center for Automated Space Science
Kentucky Space Grant and NASA EPSCoR Programs,
Western Kentucky University,
Bowling Green, KY 42101

U.S. Inan

STAR Laboratory
Stanford University

Following the first detection of an ionospheric disturbance caused by a gamma-ray burst, we searched for amplitude changes induced by gamma-ray bursts in very low frequency radio waves (VLF) propagating in the earth-ionosphere waveguide. All VLF propagation paths end in Huntsville, Alabama and the transmitters are located in Australia, Puerto Rico, Maine, Maryland, Washington State, Hawaii and Nebraska. Out of 200 bursts, which could have been detectable along 600 different propagation paths, not a single confirmed ionospheric disturbance attributable to a gamma-ray burst was found. Thus, it has been showed that the VLF approach is most likely not a viable method to detect gamma-ray bursts unless it can be demonstrated that all sea paths lead to a significant number of detections.

URC97171

RANDOM SPOTS ON CHROMOSPHERICALLY ACTIVE STARS

Joel A. Eaton and Tamara Williams

Center for Automated Space Science (CASS)

Tennessee State University, Nashville, TN 37203

S110/89

ABSTRACT

Rapidly rotating chromospherically active stars appear to be highly spotted. Heretofore we thought that the stars contained only two large spots at any one time. Eaton, Henry, and Fekel proposed that what appear to be two spots is actually the results of -20 to 30 smaller spots randomly distributed on a differentially rotating star. Initial tests of this random spot model showed that it can produce all of the characteristic photometric variations of highly spotted stars.

We have tested the random spot theory by measuring the differential rotation of the spots, the lifetimes, and the periods of migration in -20 data sets of 14 years duration in simulated model time. This time span is analogous to real stars that have been observed for about 20 years. Eaton, et al. noted that the spot lifetimes of a star with a differential rotation of $k = 0.01$ were in the range of 0.27 to 2.46 years with a median of 1.23 years. We recently calculated results of the -20 cases which show that the average differential rotation of the spots is $k = 0.012$ with a lifetime ranging from 0.26 to 3.56 years with a median of 0.55 years.

URC97172

WASHINGTON CAMP - A NEW SITE FOR TSU ASTRONOMY

Montanez A. Wade

Center for Automated Space Science (CASS)

Tennessee State University, Nashville, TN 37203

ABSTRACT

The astronomy research program at Tennessee State University began in the Center of Excellence in Information Systems in 1988 with a grant from Marshall Space Flight Center. The initial research was to expand the investigation of Dr. D. S. Hall of Vanderbilt University on the behavior of chromospherically active (CA) stars utilizing automatic photometric telescopes (APT's)located in the southern Arizonan desert at the Smithsonian Institution's Mt. Hopkins Facility (Fred Whipple Observatory). The APT's were and are operated by Fairborn Observatory, a non-profit organization. Over the years the TSU program significantly expanded and by 1996 CASS astronomers managed four APT's at the Mt. Hopkins site: Fairborn 10-in, SAO/TSU 30-in, the Vanderbilt/TSU 16-in, and the TSU/SAO 32-in. In addition to CA star research, the program now includes observation of solar duplicates to better understand the Sun-climate connection, investigation of magnetic activity in cool stars, and verification of the existence of extra-solar planets. Observing schemes are programmed in Nashville and data are retrieved from the Arizona site via the Internet. With CASS funding TSU has under construction a 2-m automatic spectroscopic telescope (AST) and a 24-in automatic imaging telescope (AIT). Anticipating this expansion, Fairborn Observatory located a 40 acre site in Washington Camp, AZ to build its own expanded facility.

This paper will present a brief history of the TSU astrophysics program including *a cursory* description of the research areas, the instrumentation utilized in data acquisition, and a description of the facilities.

Keywords: Automated astronomy, photometry, chromospherically active stars

URC97173

// // A
1172/64

LARGE SCALE DYNAMIC SYSTEMS AND CONNECTIVE STABILTY **

S. Sathananthan and Wodimu Borena *
Center for Automated Space Science (CASS)
Tennessee State University, Nashville, TN 37203, USA.

ABSTRACT

One of the foremost challenges to system theory in the present-day advanced technological world is to overcome the increasing size and complexity of the corresponding mathematical models. Since the amount of computational effort is enormous, it is simpler and more economical to decompose a large scale complex system into a number of interconnected subsystems at least for the purpose of studying stability analysis. These subsystems can be considered independent to some extent, so that some of the qualitative behavior of the corresponding subsystems can be combined with interconnection constraints to come up with the qualitative behavior of the overall large scale systems.

Besides the computational aspects of large scale dynamic systems, it is equally important to determine to what extent the complexity effects the system behavior and the role it plays in the system with large interconnected structures.

In this paper, by employing the theory of systems of integro-differential equations, a very general comparison theorem is obtained in the context of vector Lyapunov functions. Furthermore, this comparison theorem has been applied to derive sufficient conditions for the stability of the equilibrium state of the system under structural perturbations caused by the interactions among the states of the system.

** —The research reported herein is supported by the NASA research grant NCCW-0085.

*— presenter.

Keywords —Integro-Differential Equations, Stability, Interconnected Systems.

URC97174

5173/39

Photometric Monitoring of Active Galactic Nuclei in the Center for Automated Space Science: Preliminary Results

Ryan Culler, Monica Deckard, Fonsie Guilaran, **Casey** Watson, Michael Carini, Richard Gelderman

Center for Automated Space Science
Department of Physics and Astronomy
Western Kentucky University
1 Big Red Way
Bowling Green, Ky. 42101

William Neely

NF Observatory
Western New Mexico University

In this paper, we will present preliminary results of our program to photometrically monitor a set of Active Galactic Nuclei (AGN) known as Blazars. Using CCDs as N-star photometers and a technique known as aperture photometry, we can achieve close to 0.02 magnitude precision with small to midsize telescopes. Blazars are highly luminous and highly variable; studying these variations provides insight into the central engines producing the high luminosities. We report on our reduction and analysis of CCD data obtained at one of our collaborating institutions, the NF Observatory at Western New Mexico University. CCD data obtained at the Western Kentucky University 24 inch telescope will also be discussed.

N 1 1 S, 74/81

URC97175

OPTIMAL MANAGEMENT OF RENEWABLE RESOURCES IN A LARGE SCALE RESOURCE BASED COMPETITIVE SYSTEM**

S. Sathananthan and William Brown*
Center for Automated Space Science (CASS)
Tennessee State University, Nashville, TN 37203, USA.

ABSTRACT

The derivation of an optimal policy in managing natural resources has been a difficult task due to the complex nature of the biological eco-s ystem. Since the policy derived depends critically on the model used, a model containing relevant characteristics of nature is needed to obtain relevant results. Most of the mathematical models of a multi-species with harvesting studied so far have assumed the existence of an external resource for which the species compete. In this paper, we consider the growth of n-competitive species to depend not only cm an external resource but also on an (n+1)th-species. That is, there is a resource which is itself renewable.

We investigate global stability results using a suitable vector Lyapunov function. Emphasis is placed **upon** the discounted present value of the resource and the Hamilton-Jacobi equation which it satisfies. We obtained the optimal harvesting policy utilizing the Hamilton-Jacobi equation with constant effort harvesting. As a by-product, we obtained the sufficient conditions for boundedness of solutions and permanence of the system.

** —The research reported herein is supported by the NASA research grant NCCW-0085.

*— presenter.

Keywords– Stability, Optimal Control, Large-Scale Systems.

URC97176

A Study of the Planetary Nebula M57

LATONGIA GROCE

South Carolina State [University

5175-89
P. 5 3114

ABSTRACT

An examination is made of the planetary nebula M57, also known as the Ring Nebula. The presentation will begin with an overview of planetary nebulae including their role in the evolution of the Interstellar Medium (ISM). Since planetary nebulae are formed during the final stages in the life of a star, these are important objects to study. They are good laboratories for learning about how stars interact with the ISM, including how a star will enrich the surrounding medium with elements heavier than H and He. There will also be a brief discussion of the various atomic processes which are important to the nebula.

The planetary nebula M57 will be discussed next. This well known object has been the subject of many studies. What is significant about this work is that we have derived temperature and density maps of the nebula using the ions 0++, N^+ and S^+. For the first time, point-by-point variations in the electron temperature (T_e) from both the O^{++} ion and the N^+ ion can be compared. Cospatial maps of electron density (N_e) derived from S^+ are also presented. These quantities are important in order to calculate the chemical abundances of this object and to better understand how the ISM will evolve over time.

These maps were generated using CCD images taken at the 2.12 meter telescope at San Pedro Matir in Baja, Mexico. The images were calibrated on a SPARC 20 workstation using the well known image processing software, IRAF. The standard reduction steps were followed, including bias and dark subtraction, flat field corrections and cosmic ray removal. These steps will be discussed briefly.

URC97177

5/76/ ...

SURFACE PHONONS AND SURFACE PLASMONS IN QUANTUM DOTS AND METAL COLLOIDS

D. O. Henderson, R. Mu, A. Ueda, Y. S. Tung
Chemical Physics Laboratory and NASA Center for Photonic Materials and Devices
Physics Department, Fisk University, Nashville, TN 37208

Jane G. Zhu, C. W', White and R. Zuhr
Oak Ridge National Laboratory, Solid State Division
Oak Ridge TN 37831-6057

ABSTRACT

Nanophase materials have attracted considerable attention from the scientific and technological communities. The intense interest in this class of materials originates from the fact that reduction of matter to the nanometer scale often results in a manifestation of material properties that are unique from the bulk. The technological implication of nanophase materials has already had an impact in the areas of catalysis, optical devices for telecommunications, lasers based on quantum well structures, high density optical memory, optical computing, and ultrafast all-optical switching devices. Nanophase materials are also ideal candidates for investigating cross-over phenomena, whereby a metal cluster, for example, undergoes a transition from having physical properties that can be described by molecular physics to those which can be understood in terms of band structure theory of the bulk. Nanophase materials also open the opportunity to study the basic physics and chemistry that govern the transformation of very small clusters to subcritical nuclei to critical nuclei and consequently to the crystal growth process itself.

In the course of the transition, quantum size effects are often observed and are expressed in the optical properties of the nanophase materials. As an example, semiconductor nanocrystals with particle sizes smaller than their excitonic radii (quantum dots, QD's) show blueshifts in their bands gaps as a result of quantum confinement. The phonon spectra of QD's exhibit surface modes that arise from a break-down of the cyclic boundary conditions for the bulk material and lead to a resonance that falls between the bulk longitudinal and transverse optical phonon frequencies. Metal nanocrystals, on the other hand, show collective-ti&d&, surface plasmons, in the visible to the ultraviolet part of the spectrum and are responsible for some of the beautiful colors of metal doped glasses. Such optical properties of metal nanocrystals and quantum dots can be used as probes to interrogate cross-over phenomena to gain an understanding of how the properties of nanophase materials evolve as a function of size. Along this line, we have investigated the optical spectra of gold nanocrystals embedded in dielectric matrices and report on the development of the surface plasmon absorption (530 nm) for different annealing environments. Surface phonon spectra for GaAs nanocrystals buried in sapphire and silica matrices are also reported are and shown to agree with Frölichs theory of surface phonons

Corresponding. Author: D. O. Henderson
Chemical Physics Laboratory
Department of Physics
Fisk University
Nashville, TN 37208
E-mail: hendersn@dubois.fisk.edu